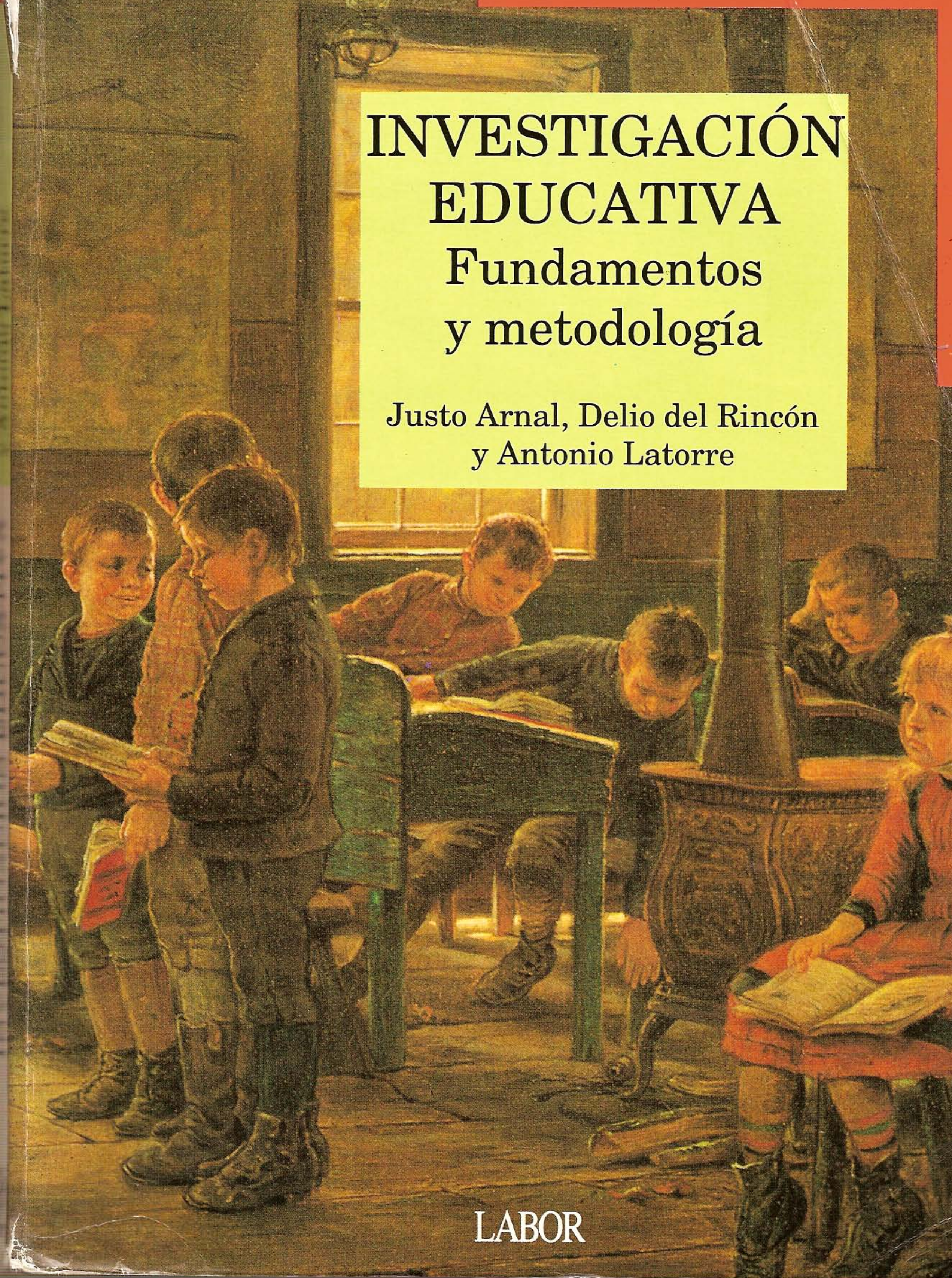
LABOR UNIVERSITARIA

INVESTIGACIÓN EDUCATIVA
Fundamentos y metodología

Justo Arnal, Delio del Rincón
y Antonio Latorre



LABOR

1.

Investigación educativa

Fundamentos y metodologías

Justo Arnal,
Delio del Rincón,
Antonio Latorre

Investigación educativa

Fundamentos y metodologías

EDITORIAL LABOR, S.A.

Cubierta de Jordi Vives

Primera edición: 1992

ISBN: 84-335-3725-3
Depósito legal: B. 29.974-1992
Printed in Spain - Impreso en España
Impreso en Gráficas 92 *Cl* Torrasa, 108-E Sant Adrià de Besos

Índice

II. METODOLOGÍAS DE INVESTIGACIÓN EDUCA TIV A DE PERSPECTIVA EMPÍRICO-ANALÍTICA

III. METODOLOGÍAS DE INVESTIGACIÓN EDUCATIVA DE PERSPECTIVA HUMANÍSTICO-INTERPRETATIV A

Presentación

Las últimas reformas educativas abogan por una mayor *profesionalización* de los docentes y destacan la capacidad investigadora de los profesionales de la educación como uno de los elementos clave e imprescindible para responder al reto de mejorar la calidad de la enseñanza.

Esta obra surge con la finalidad de hacer posible que los profesionales de la educación afronten el "reto" de mejorar la educación a través de la actividad investigadora. Pretende ofrecer a quienes se forman en el quehacer investigador los elementos fundamentales y las herramientas de trabajo necesarias para poder llevar a cabo investigación educativa. La finalidad de la obra en sí es doble: por una parte llenar el vacío existente en nuestro medio cultural hispanoamericano de obras de esta naturaleza, y por otra, presentar una obra *sobre investigación educativa* de carácter básico e introductorio que a la vez que ofrece una panorámica general de la investigación recoge las nuevas tendencias metodológicas que han surgido en las últimas décadas en este ámbito. *Investigación educativa. Fundamentos y metodologías* pretende ser, en definitiva, una obra-guía que proporcione a los profesionales de la educación y estudiantes en formación los conocimientos y recursos básicos y necesarios para adentrarse en la realización de investigaciones.

El texto se ha estructurado en tomo a dos grandes bloques temáticos: *Fundamentos de investigación* y *metodologías.* En el bloque de fundamentos de investigación, que abarca los tres primeros capítulos, se describen los elementos básicos que conforman el quehacer investigador. Su propósito es establecer las bases en las que se asienta la actividad investigadora. En el primer capítulo se introducen algunos conceptos básicos de investigación, como conocimiento científico, ciencia, teoría, metodología científica, método científico e investigación científica, que constituyen los fundamentos de la investigación en general. El capítulo segundo, de carácter más conceptual, se dedica a estudiar la naturaleza de la investigación educativa, considerando aquellos aspectos que la definen y delimitan, como su perspectiva histórica, sus diversas concepciones y características, los diferentes paradigmas en los que se apoya, la amplia gama de modalidades de investigación existentes y los aspectos referidos a sus límites y deontología. En el capítulo tercero se describe el tema clave del proceso general de investigación, sin entrar en los procesos específicos de cada metodología, de ahí la denominación de proceso general. Se señalan y describen de una manera muy sucinta los distintos momentos o fases del proceso de investigación, desarrollando con más detalle las fases deductivas del proceso: el planteamiento del problema, la revisión de la literatura, las hipótesis y variables, la muestra y la introducción a las perspectivas metodológicas.

El segundo bloque temático se dedica al estudio de las distintas metodologías de investigación. Éstas se han agrupado en tres grandes perspectivas metodológicas: la empírico-analítica, la humanístico-interpretativa y la orientada a la práctica educativa. Las dos primeras metodologías poseen aspectos epistemológicos y metodológicos propios, y la tercera, que los toma de las anteriores, dispone de procesos y diseños de investigación propios.

Los capítulos cuarto al séptimo estudian y describen las metodologías de orientación empírico-analítica, con sus correspondientes modalidades de investigación experimental, cuasi-experimental y ex post facto. Dentro de cada metodología se señalan sus rasgos principales, sus modalidades y diseños más representativos. Estas metodologías se tratan con cierta extensión debido al desarrollo e implantación que tienen en algunos ámbitos educativos.

En los capítulos octavo al décimo se estudia las metodologías de orientación humanístico-interpretativa. Se señalan los rasgos que definen estas metodologías y se exponen de forma resumida algunas de las nuevas tendencias de investigación que han emergido en las últimas décadas, como la fenomenología, el interaccionismo simbólico, la etnometodología, el estudio de casos o la etnografía. En la actualidad estamos asistiendo a un gran auge de estas metodologías en el campo de la educación.

Los capítulos once y doce recogen y describen las metodologías que hemos denominado orientadas a la práctica educativa. Son metodologías que se singularizan por su carácter de aplicación a la política educativa, es decir, al cambio y la mejora de la práctica educativa. En la actualidad están teniendo una gran incidencia en educación y se espera que contribuyan a mejorar la práctica educativa. Dentro de esta perspectiva metodológica se consideran la investigación evaluativa, haciendo especial hincapié en el diseño respondente, y la investigación acción con sus modalidades de investigación participativa y colaborativa. La investigación acción está llamada a ocupar un lugar importante y a constituirse en la metodología más relevante para los prácticos de la enseñanza.

La estructura del libro, su exposición teórico-práctica, la sistematización de los contenidos, su lenguaje asequible y comprensible, la profusión de tablas y figuras, etc., son algunos de los elementos que proporcionan a la obra un carácter eminentemente didáctico con el fin de que resulte útil para el aprendizaje de los lectores.

Como se deduce de lo señalado más arriba, la obra va dirigida a los estudiantes universitarios, en especial a los que cursan estudios de pedagogía, psicopedagogía, educador social y magisterio y, en general, a los profesionales de la educación que deseen formarse y actualizarse en el conocimiento de los procesos y métodos propios de la investigación educativa.

Por último, no queremos terminar sin agradecer los comentarios y críticas que sobre los contenidos y formato de la obra nos han hecho llegar nuestros colegas de los respectivos departamentos, y animamos a los lectores a continuar en esta línea aportando sus puntos de vista, sugerencias y críticas sobre cualquier aspecto de la obra que crean oportuno. Tengan por seguro que sus opiniones y sugerencias serán bien recibidas.

Barcelona, julio de 1992

PARTE I

FUNDAMENTOS DE LA INVESTIGACIÓN EDUCATIVA

CAPÍTULO 1

Nociones básicas sobre investigación

Hacer preguntas es una actividad específicamente humana. A lo largo de la historia el hombre ha sido siempre un ser preocupado por entender y desentrañar el mundo que le rodea, por penetrar en sus relaciones y leyes, por orientarse hacia el futuro y descubrir el posible sentido de las cosas que existen a su alrededor, buscando respuestas a sus interrogantes.

Pueblos de la Antigüedad, como el griego, intentaron comprender el mundo y sus fenómenos mediante la razón. En la actualidad, sin embargo, el hombre utiliza diversas fuentes de conocimiento como la experiencia, los expertos o el razonamiento, ya sea deductivo, que permite pasar de lo general a lo particular, o inductivo, que partiendo de los datos particulares llega a generalizaciones.

1.1 TIPOS DE CONOCIMIENTO

Según las fuentes utilizadas y las características del objeto que intentamos conocer podemos hablar de conocimiento vulgar, práctico, artístico, religioso, técnico, filosófico o científico. Estas formas de conocimiento pueden resultar muy valiosas en una situación concreta; sin embargo, el conocimiento científico es el más vinculado a la investigación. En cuanto a las demás formas de conocimiento, si bien necesarias y útiles, no son suficientes para obtener conocimiento científico.

a) *Conocimiento vulgar*. La experiencia revela que la vía ordinaria que sigue el hombre para resolver sus problemas suele basarse en el conocimiento vulgar, por ser una forma de conocimiento práctico, que se transmite directamente de unos a otros y se manifiesta, en parte, en la cultura popular. Es un pensar espontáneo que preside la vida cotidiana. De alguna manera, el saber vulgar es propio del sentido común, y éste se concibe como un campo de conocimiento propio de la comunidad que permanece sin articular y sin conformar, pero que es de importancia inmediata para entorno s más amplios de la experiencia y práctica humanas. El sentido común se forma lenta y cuidadosamente, y se estima como condición de la sociabilidad y de la comunidad en los aspectos prácticos de la vida.

El conocimiento vulgar (Wartofsky, 1983, 91) no es explícitamente sistemático ni

crítico, por estar basado también en la autoridad o en la tradición, de modo que ninguna de sus partes atañe a todas las demás, ni existe intento consciente por considerarlo como cuerpo consistente de conocimientos. Sin embargo, en general, es un conocimiento completo y se encuentra listo para su utilización inmediata, como corresponde a la propiedad común de la cultura, por constituir un conjunto de previsiones certeras y referidas a lo que todo el mundo debiera saber de las actividades cotidianas y básicas de la vida diaria. Las funciones que desempeña, posibilitando el trabajo ordinario y la vida social, son de importancia inapreciable. Es probable que esta forma de conocimiento, a pesar de su fragilidad, resulte muy útil y sea la más frecuente en la práctica educativa cotidiana. Pero la crítica y la sistematización en el ejercicio del pensamiento humano es crucial. De ahí que una diferencia importante entre ciencia y sentido común resida en· que la ciencia intenta ser consciente y deliberadamente más crítica y organizada. En este sentido, el saber filosófico y el científico son formas de conocimiento más desarrolladas.

b) *Conocimiento filosófico.* Trasciende la percepción inmediata para buscar el porqué de los fenómenos y se basa fundamentalmente en la reflexión sistemática para descubrir y explicar. Sin embargo, en el ámbito educativo, este tipo de conocimiento está expuesto a un doble peligro: la imprecisión y la falta de contrastación con la realidad en aquellos contenidos que podrían ser más susceptibles de análisis empírico.

c) *Conocimiento científico.* Es uno de los modos posibles del conocimiento humano. No es el único capaz de ofrecer respuestas a nuestros interrogantes; sin embargo, es el más útil y desarrollado, y, por hallarse muy vinculado al proceso de investigación, lo abordaremos más detalladamente.

1.2 CONOCIMIENTO CIENTÍFICO

Los límites entre el conocimiento científico y el vulgar no están claros, pues si bien existe una estrecha relación entre ellos, también se da una diferencia de grado. En cuanto a las semejanzas entre ambas modalidades de conocimiento, hay que destacar las pretensiones de racionalidad y objetividad; esta afinidad puede explicar que para algunos hombres de ciencia como Popper (1977, 19-22) el conocimiento científico sea el desarrollo del sentido común. Ahora bien, el conocimiento científico es más afinado que el vulgar por aventurar conjeturas (Bunge, 1985) y por criticar las explicaciones logradas a través del sentido común (Sarramona, 1980,3), por su extensión, claridad y precisión, por centrarse más en el porqué de las cosas y por su carácter sistemático y seguridad comprobada (Pérez Juste, 1981, 19). De ahí que, en ocasiones, el científico pueda elevarse por encima del saber vulgar, del sentido común, en la búsqueda progresiva de mejores soluciones a los problemas. La historia ofrece situaciones y personas, como Galileo, que incluso mantuvieron posturas opuestas al sentido común de sus contemporáneos.

Por otro lado, el conocimiento científico queda caracterizado por el método que utiliza y, por ello, puede definirse como todo conocimiento adquirido a través del método científico. Quizá el criterio clave de distinción entre ambos conocimientos esté en el tipo de explicación en que se apoyan. Las características del conocimiento

científico han sido ampliamente comentadas por autores como Bayés (1974), Arnau (1978), McGuigan (1983) y Kerlinger (1985), quienes 10 caracterizan como:

a) *Objetivo.* Un conocimiento es objetivo cuando se corresponde con la realidad del objeto y lo describe o explica tal cual es, y no como nosotros desearíamos que fuera; consiste en tratar de encontrar la realidad del objeto o fenómeno estudiado, elaborando proposiciones que reflejen sus cualidades.

El conocimiento científico es independiente de las opiniones o preferencias individuales, es imparcial y comprobable mediante la replicación. No tiene prejuicios e intenta ver las cosas tal como son «realmente». La desvinculación del investigador respecto a lo investigado contribuye a que los resultados de la investigación sean independientes del investigador y permite comprobarlos mediante su replicación y su coherencia. Autores como Kerlinger (1985) señalan la necesidad de un acuerdo entre jueces expertos respecto a lo que se ha de observar, lo que se ha de hacer o lo que se ha hecho en investigación; Popper (1977, 46) señala que la ciencia se ocupa de enunciados objetivos que han de poder ser contrastados intersubjetivamente.

La objetividad se considera posible gracias al contraste intersubjetivo, a la fiabilidad de la percepción y del razonamiento (Brown y Gisselli, 1969; Popper, 1988) ya la comprobabilidad de las afirmaciones. Ahora bien, de ello no se deduce una única lectura e interpretación de los hechos o fenómenos, pudiendo ser aceptables varias interpretaciones si cada una de ellas ha sido contrastada. Así, el conocimiento científico es falible, inexacto y provisional, pues verdades tenidas por objetivas.. definitivas y demostradas en un momento histórico son descartadas o bien sustituidas en otro. De ahí que, siguiendo a Popper (1977, 261), la objetividad deba considerarse como característica que hace de todo enunciado científico algo «provisional para sIempre».

b) *Fáctico.* La fuente de información y de respuesta a los problemas es la experiencia, es decir, hechos o fenómenos de la realidad externos al investigador. Autores como De la Orden (1981), Kerlinger (1985), Piaget (1982) y otros resaltan esta peculiaridad del conocimiento científico, considerando que lo que distingue a este conocimiento es la base empírica y la experiencia. Esto no significa, sin embargo, que exista restricción por parte de la ciencia al utilizar tal término, puesto que no se refiere necesariamente a hechos o fenómenos perceptibles u observables de manera directa.

c) *Racional.* Otra característica del conocimiento científico es el hecho de que la ciencia utiliza la razón como vía esencial para llegar a sus resultados. Esta característica permite, según Bunge (1985), la sistematización coherente de enunciados fundados o contrastables, y el logro de una teoría o un conjunto sistemático y racional de ideas sobre la realidad de que se trate.

Ante el carácter racional de la explicación ofrecida por el método científico, parte de las acciones y fenómenos humanos no tienen una justificación racional, y no se pueden comprender fuera de su situación y contexto. Algunos especialistas actuales (Blasi, 1980, y Locke, 1983) han resaltado la falta de una metodología adecuada para estudiar la estructura y organización de las acciones humanas que no siempre tienen una base lógica y causal (Packer, 1985). En este sentido, como ya apuntamos,

deberíamos admitir otros tipos de «racionalidad» que no se excluyen del método científico.

Nuevas aportaciones, tanto epistemológicas como pedagógicas, hacen más amplio y flexible este término, no reduciéndolo a las pautas de la lógica inferencia!. En él tienen cabida otras vías «racionales» (Polanyi, 1958; Morin, 1984), que pueden ser prometedoras en el ámbito educativo.

d) *Contrastable.* La contrastabilidad permite una mayor fiabilidad del conocimiento (Bunge, 1981; Popper, 1988) al ser comprobado por distintas personas y en 'circunstancias variadas. Es un intento de conjunción entre lo fáctico y lo racional que implica la puesta en marcha de diversas técnicas y procedimientos en las investigaciones científicas. Este término incluye la contrastabilidad teórica e indirecta (Bunge, 1981), que actúa a nivel más complejo y profundo que el de la realidad «directamente observable», utilizada en el caso de las contrastaciones directas. De esta manera el conocimiento científico debe someterse a prueba, contrastándose intersubjetivamente a través de la experiencia a la luz de los conocimientos objetivos por medio de nuestro ingenio y conocimiento en una discusión y examen críticos.

e) *Sistemático.* Es un conocimiento ordenado, consistente y coherente en sus elementos, una totalidad interrelacionada e integrada en un sistema. Un conocimiento aislado no puede considerarse científico. Bochenski (1981,30) pone de relieve que «no todo el que posee conocimiento de algún dominio del saber posee ciencia de él, sino solamente aquel que ha penetrado sistemáticamente y que, además de los detalles, conoce las conexiones de los contenidos». En este sentido, Kerlinger (1985) resalta que el científico busca consciente y sistemáticamente relaciones. Mediante la sistematización se organizan, se estructuran y armonizan diferentes conocimientos, para tratar de entender la realidad de la manera más completa posible (Yurén, 1980)

La sistematización conlleva dificultades originadas por la gran cantidad de variables que guardan estrecha interrelación, pudiendo haber tantas teorías y explicaciones como personas las formulen (Blalock, 1985). Pero también existe la posibilidad de contrastar las distintas formulaciones teóricas.

f) *Metódico.* El conocimiento científico es fruto de una metodología rigurosa. Se obtiene mediante la aplicación de planes elaborados cuidadosamente para dar respuestas a preguntas o problemas. Su condición de científico se apoya en la fiabilidad de los procedimientos y estrategias utilizadas para su obtención.

g) *Comunicable.* Expresado en lenguaje apropiado y preciso, en términos de significación inequívoca reconocidos y aceptados por la comunidad científica. La utilización de un lenguaje claro y preciso (Nagel, 1979) facilita la comunicabilidad del contenido, favoreciendo así el carácter autocorrectivo del proceder científico (Kuhn, 1984; Popper, 1988). La confusión terminológica hace difícil una comunicación efectiva, se hace imprescindible un lenguaje claro y preciso que permita un intercambio de información más efectivo.

h) *Analítico.* La forma de proceder «analítica» para obtener el conocimiento científico obliga a seleccionar variables o cuestiones que «rompen» la unidad,

complejidad-y globalidad de los fenómenos humanos, actuando a distintos niveles y con diversos grados de globalización. La contrastabilidad y la objetividad se ven, facilitadas por este procedimiento analítico que consiste en seccionar la realidad para poder abordarla con mayores garantías (López Cano, 1978). Al faltar un patrón universal de medida para decidir a qué niveles fragmentar la realidad, un mismo objeto de estudio puede ser entendido y abordado de maneras distintas y con grados de molecularidad diversos (Guba, 1982; Patton, 1984; Cook y Reichardt, 1986), posibilitando, a su vez, distintas decisiones en cuanto a su nivel de generalización.

En conjunto, las características descritas diferencian el conocimiento científico de otros tipos de conocimiento. Sin embargo, algunas características tienden a predominar más en determinadas disciplinas. Así, el aspecto fáctico es importante en disciplinas artísticas e incluso en el conocimiento popular (Fletcher, 1984). La tradición empirista ha primado los aspectos fácticos, analíticos y la contrastación directa. En cambio, según el punto de vista racionalista, la teorización o sistematización constituye un elemento esencial en la generación del conocimiento, que se valida a través del conocimiento lógico-matemático (Packer, 1985).

t,,:3 CIENCIA: ESTRUCTURA Y FUNCIONES

La ciencia es una vasta empresa que exige un gran esfuerzo humano con el objetivo de adquirir conocimientos válidos sobre la realidad. Es una de las actividades que el hombre realiza como un conjunto de acciones encaminadas y dirigidas a obtener conocimientos contrastrab1es sobre los hechos que nos rodean.

Este apartado describe qué se entiende por ciencia, sus elementos, estructura y funciones. Posteriormente nos centraremos en el método científico que nos permite llegar a la ciencia.

Para clarificar mejor la estructura y funciones de la ciencia, previamente delimitaremos sus términos básicos: hechos, fenómenos y datos.

Hechos, fenómenos y datos

En general, un hecho es todo aquello que se sabe o se supone, con algún fundamento, que pertenece a la realidad. Por ejemplo: la vocalización de una palabra, un aula, un profesor explicando, la respuesta dada a un test. Como puede apreciarse, existe una gran variedad de *hechos,* que a su vez pueden considerarse como acontecimientos, procesos y sistemas concretos (Yurén, 1980).

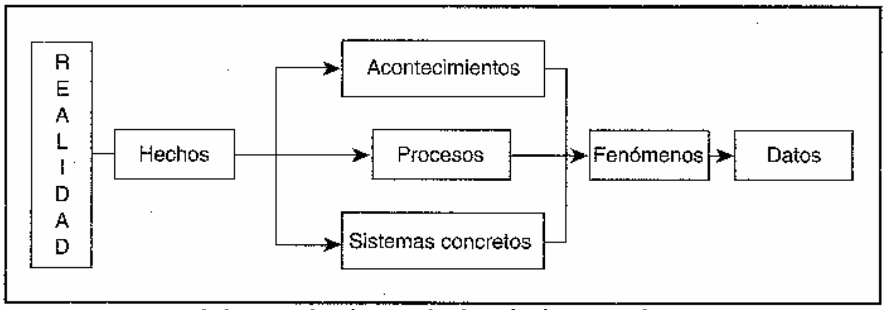
Los hechos que tienen lugar en el espacio y en el tiempo, como pueden ser un grito, un golpe, una pregunta o una respuesta, se denominan *acontecimientos.* Cuando varios acontecimientos constituyen una secuencia temporalmente ordenada, de tal manera que cada acontecimiento implica, afecta o permite los siguientes, estamos ante un *proceso.* Así, en un proceso didáctico podrían estar implicados acontecimientos como la explicación del profesor y las anotaciones y respuestas del alumno. Ahora bien, en sentido estricto, podemos considerar que la mayoría de los acontecimientos están constituidos, a su vez, por procesos. Por ejemplo, un profesor explicando puede considerarse como un acontecimiento y, a su vez, como un proceso de acontecimientos

si consideramos que pronuncia palabras, las escribe en la pizarra y son percibidas por los alumnos.

Por último, cuando un hecho está constituido por partes identificables y estructuradas, formando un todo unitario, estamos ante un *sistema concreto,* como puede ser un aula, un taller de pintura, un audiovisual o un claustro de profesores. Para conocer y dominar la realidad que le rodea, el ser humano percibe y capta, por medio de *sus* sentidos, una gran variedad de hechos que constituyen acontecimientos, procesos y sistemas concretos.

Cuando son percibidos por el investigador, los hechos observables suelen denominarse *fenómenos;* es decir, son captados por el observador y percibidos a través de los sentidos como fenómenos; en consecuencia, el fenómeno designa un hecho percibido.

El conocimiento que se extrae de los fenómenos lo denominamos *información;* ésta se transmite en forma de *datos,* que es una información fijada o codificada por el investigador (fig. 1.1).



e)

FIG. 1.1 Esquema de la interrelación entre hechos, fenómenos y datos

Cualquier información, por pequeña o fragmentaria que sea, puede considerarse como un dato. Por tanto, el dato encierra enunciados, afirmaciones o negaciones que expresan aspectos o características concretas de la realidad (Sierra Bravo, 1984, 121). Por ejemplo, el sujeto X es niño, tiene 10 años y obtuvo una puntuación de 100 en un test de inteligencia.

Relaciones entre fenómenos

Sólo una pequeña parte de los hechos que ocurren son observables; a partir de estos fenómenos el científico prosigue la búsqueda de conocimiento, utilizándolos como evidencias que confirman o sugieren la existencia de otros hechos cuyas manifestaciones hemos observado, o bien de conexiones y correspondencias entre estos hechos. Así, dos niños peleándose en el patio de juegos es un hecho que puede constituir un fenómeno para el educador. La agresión puede ser una manifestación de otro hecho subyacente como es la privación de una gratificación esperada. Si así fuera, existiría una relación entre la privación y la agresión. En consecuencia, una conexión o

correspondencia entre fenómenos suele denominarse relación y se expresará mediante proposiciones. Algunos ejemplos de relaciones podrían ser:

a) Juan no dejó el juguete a Rosa y ésta le agredió.
b) Con incentivos positivos, aumentan las respuestas correctas.
c) La falta de actividad adecuada origina indisciplina.
d) El calor dilata los metales.

Como puede apreciarse en los ejemplos, existe una relación entre los fenómenos que aparecen *subrayados.* Ahora bien, entre los mismos fenómenos pueden existir otras formas distintas de relación. Por ejemplo, entre los fenómenos «horas de estudio dedicadas», «aumento de respuestas correctas» y «recibir los elogios del profesor» podrían existir relaciones como las siguientes:

f) Las horas de estudio dedicadas hicieron que aumentaran las respuestas correctas, y por ello, la clase recibió elogios del profesor.
±) Debido a que la clase recibió los elogios del profesor los alumnos dedicaron horas de estudio y esto hizo que las respuestas correctas aumentaran.
g) Las respuestas correctas aumentaron, y por ello, el profesor elogió a la clase y los alumnos dedicaron horas de estudio.

En los ejemplos mencionados existe una conexión entre los fenómenos implicados. Los fenómenos son -los mismos, pero varía la secuencia temporal y, presumiblemente, la relación de causalidad entre ellos. A la forma en que se relacionan las partes de un todo se la denomina estructura. Ésta cambia cuando varía la relación entre las partes, ya sean físicas, o bien cualidades, aspectos o abstracciones que llamaremos elementos. Así, en la figura 1.2 aparecen tres estructuras formadas cada una de ellas por un conjunto de cuatro letras. Las letras no cambian, lo que varía es la relación entre ellas. La estructura 1 es rectangular; la 2 es triangular, y la 3, lineal. Cada estructura tiene cuatro letras bien diferenciadas, pero siempre son las mismas. Sin embargo, en la estructura 1 las letras A y B mantienen la relación «estar encima de» las letras C D, mientras que en la estructura 3 mantienen una relación de «estar al lado de» (fig. 1.2).

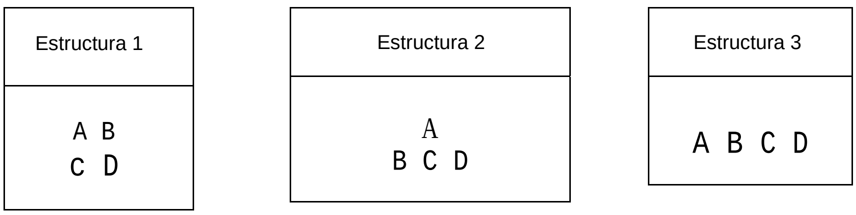


FIG. 1.2 Ejemplos de estructuras

Dado que el investigador se interesa fundamentalmente por determinado tipo de relaciones, nos detendremos brevemente en su descripción.

A) CLASES DE RELACIONES

Las relaciones entre los fenómenos pueden ser de distinta naturaleza. Así, la relación *a*, «Juan no dejó el juguete a Rosa y ésta le agredió», expresa una conexión entre fenómenos particulares, no es generalizable a otros sujetos. No dice que todos los niños que se niegan a dejar sus juguetes son agredidos. Además, la relación *a* podría haberse presentado de otra manera; es decir, no era forzoso que Rosa agrediera a Juan, sino que Rosa podría haberse marchado llorando; por ello, no es una relación necesaria. De la misma manera, dicha relación no es constante, no ocurre siempre, es decir, todos los niños que se niegan a dejar sus juguetes no siempre son agredidos. A veces sucede así y a veces no.

En cambio, las relaciones *b* y e son más generales. Por ejemplo, los incentivos positivos, en general, y no un tipo especial de incentivos, tienden a aumentar las respuestas correctas. Estas relaciones son más necesarias y constantes que la relación *a*. Como podrá observarse, estas características aparecen todavía en mayor grado en la relación *d*, ya que todos los metales en general se dilatarán con el calor necesariamente; no puede suceder de otra manera; es una relación necesaria y se dilatarán siempre (relación constante). De ahí que una relación será constante en la medida que la consideremos general y necesaria. Cuando las relaciones sean generales, necesarias y constantes se constituirán en leyes. Si en una estructura consideramos lo permanente de la r-elación, independientemente de los cambios que pueden experimentar las partes, aspectos o propiedades de los fenómenos implicados, entonces consideramos una relación constante que se denomina *ley* (Yurén, 1980).

B) LEYES: DESCUBRIMIENTO Y CONTRASTRACIÓN

La ciencia se interesa por las relaciones entre los fenómenos y, en la medida de lo posible, tiende a configurarse en base a relaciones progresivamente más generales, necesarias y constantes. Dichas relaciones se irán aceptando después de ser debidamente contrastadas en la realidad. Se considerarán leyes cuando sean relaciones constantes e invariables.

El investigador puede estar interesado en descubrir relaciones y leyes desconocidas a partir de unos datos, o bien en comprobar la existencia de una ley previa.

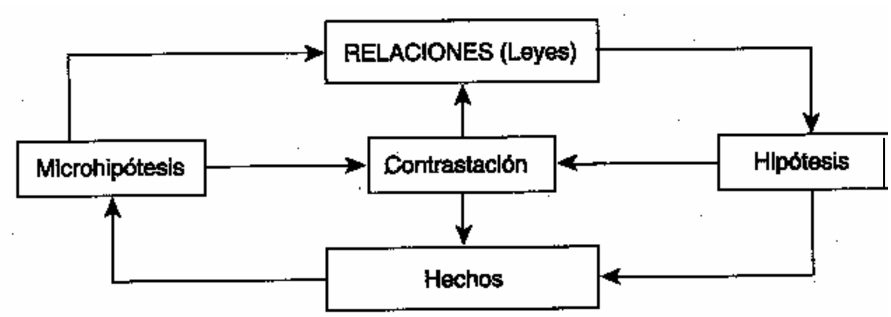


FIG. 1.3 Descubrimiento y contrastación de relaciones

recurriendo a datos recogidos al observar los fenómenos implicados. A partir de ellos el investigador reflexiona, los estudia, analiza y ordena para descubrir si existe relación entre ellos, o para comprobar la viabilidad de una relación previamente establecida. Tanto para descubrir como para contrastar dichas relaciones hay que recurrir a lo que denominamos *investigación científica.* En general, la investigación se realiza para aportar información en tomo a un problema planteado. Según la naturaleza del problema el proceso de investigación se orientará a descubrir relaciones o microhipótesis, o bien a contrastar relaciones previamente establecidas o hipótesis (fig. 1.3).

Tanto las hipótesis como las microhipótesis constituyen una respuesta provisional o explicación del problema planteado en forma de relación entre fenómenos. En consecuencia, esta respuesta provisional, en sentido amplio, puede denominarse genéricamente *hipótesis.*

e) FUNCIONES DE LAS RELACIONES

Las funciones fundamentales de las relaciones son las inherentes al conocimiento científico: explicar y predecir. Cuanto más general, necesaria y constante sea una relación, mejor podrá hacer predicciones en tomo a los fenómenos implicados. Por ejemplo, según la relación *b* podemos predecir que aumentarán las respuestas correctas de un grupo de alumnos si se les aplican incentivos positivos. De la misma manera, a partir de la relación *d* podemos predecir con mayor seguridad que un trozo de hierro se dilatará al calentado.

Cuando una relación como las anteriores implica una secuencia temporal entre fen6menos, y se constata que un fenómeno aparece sólo cuando está presente otro fenómeno denominado *antecedente,* puede afirmarse que dicha relación permite' explicar el fenómeno consecuente basándose en la relación que mantienen ambos. Así, la relación *b* contiene:

- un fenómeno antecedente: «Aplicación de incentivos positivos»;
- un fenómeno consecuente: «Aumento de respuestas correctas».

Si los datos recogidos indican que el aumento de respuestas correctas sólo está presente cuando se administran incentivos positivos podemos decir que la relación *b* permite explicar el fenómeno del aumento de las respuestas correctas en un grupo determinado de alumnos.

Para poder explicar y predecir los fenómenos es necesario elaborar un conjunto o sistema de relaciones, contrastadas mediante datos empíricos. Este sistema de relaciones constituye lo que denominaremos *teoría.*

Teoría científica

Después de haber delimitado el conocimiento científico con respecto a otros modos de conocer, conceptualizaremos la ciencia, describiendo brevemente cómo se estructura internamente y mediante qué procesos se elabora. Para una mayor ampliación y clarificación de las tendencias actuales en la manera de concebir la ciencia pueden

consultarse como referencia directa las obras de algunos filósofos de la ciencia (Yurén, 1980; Lakatos, 1983; Hampel, 1987; Popper, 1988, y Bunge, 1985a y 1988) y las aportaciones de otros autores (Sierra Bravo, 1984; Rosel, 1986, y Dendaluce, 1988) sobre el mismo tema.

A) LA CIENCIA

La palabra *ciencia* deriva del latín *scientia,* que tiene un sentido amplio y significa «conocimiento, doctrina, erudición o práctica». En su acepción general, por tanto, ciencia equivale a toda clase de saber. En el curso del tiempo el concepto de ciencia se precisó, llegando a significar un conjunto de conocimientos sistemáticos sobre una disciplina o materia académica. Se puede decir que la disciplina por antonomasia en la Edad Media era la filosofía. En la actualidad, con el desarrollo de nuevos procedimientos de adquirir conocimiento basados en la observación y experimentación, el concepto de ciencia ha quedado reservado para el conocimiento científico. En este sentido se entiende por ciencia *el conjunto organizado de conocimientos sobre la realidad y obtenidos mediante el método científico (Sierra* Bravo, 1984).

En cierto modo, la ciencia trata de paliar el posible énfasis concedido al método reflexivo para desarrollar los argumentos filosóficos, evitando que las teorías sean puras abstracciones conceptuales sin un contenido empírico (Rosel, 1986). El conocimiento científico intenta subsanar el posible alejamiento de aquellos aspectos susceptibles de ser contrastados con la realidad. Para ello la ciencia conjuga el sentido común y el pensamiento reflexivo con la contrastación empírica, originando el conocimiento científico.

Desde un planteamiento globalizador (Dendaluce, 1988), la mayoría de los autores definen la ciencia en función de los componentes contenido, método y producto, y la conciben como un *modo de conocimiento riguroso, metódico y sistemático que pretende optimizar la información disponible en tomo a problemas de origen teórico y/o práctico.*

Clasificación de las ciencias. Si se tiene en cuenta que el conocimiento se refiere siempre a un contenido, las ciencias pueden clasificarse *empíricas* y *formales,* según que su contenido haga o no referencia a los hechos. Las ciencias denominadas *formales* se ocupan de estudiar relaciones, pero sin referirlas a los hechos. El contenido de las ciencias formales son entidades lógicas o matemáticas. Ejemplos de este tipo de ciencias son la filosofía y las matemáticas. *Las ciencias empíricas o factuales* se ocupan de estudiar los hechos y sus relaciones, pero siempre referidas a los hechos (Yurén, 1980). El contenido de las ciencias empíricas son los hechos percibido s como fenómenos a través de la experiencia. Así, ciencias como la física: la química o la biología estudian los hechos naturales. Otras ciencias empíricas como la sociología, la economía, la política, la antropología y el derecho se ocupan de los hechos sociales y, especialmente, algunas ciencias de la educación se ocupan de los hechos educativos.

Funciones de la ciencia. Según las aportaciones de reconocidos especialistas (Kerlinger, 1985; Tejedor, *1985a;* Bunge, *1985a;* Rosel, 1986, y Keeves, 1988), pueden

extraerse algunas funciones de la ciencia que están implicadas en las teorías científicas: *comprensión, explicación, predicción* y *control.* La ciencia también puede orientar la toma de decisiones y los procesos de cambio o transformación de la realidad. Sin embargo, los científicos no suelen conceder la misma importancia a todas las funciones mencionadas.

Para autores como Kerlinger (1985,6), el objetivo fundamental de la ciencia es la explicación de los fenómenos. La explicación, como veremos más adelante, es un componente esencial de las teorías. Pero, en general, la mayoría de los autores están de acuerdo en afirmar que explicación y comprensión son esenciales, por ser decisivas para la toma de decisiones y para modificar y predecir, en la medida de lo posible, t:L curso de los fenómenos.

La diferenciación entre explicación y comprensión la estableció Dilthey al señalar que el objetivo de las ciencias de la naturaleza es «explicar», sobre todo basándose en las semejanzas y regularidades entre los fenómenos, y el de las ciencias humanas es «comprender», atendiendo también a las diferencias entre los sujetos. De ahí que el reto de las ciencias humanas sea conjugar adecuadamente, y en el grado que sea posible, la sistematización explicativa de los sujetos y de las situaciones, con la identidad propia de cada sujeto y cada contexto. En términos generales, las funciones de la ciencia se alcanzan a través de las teorías.

B) LAS TEORÍAS

Las ciencias guardan un orden en sus conocimientos. Toda ciencia necesita estructurar sus conocimientos, relacionarlos y configurar una estructura o sistema que se denomina *teoría.* La conexión, la estructura y el orden que guardan los conocimientos constituyen su aspecto f9rmal (Yurén, 1980). Las ciencias empíricas tienen una estructura o sistema que se configura mediante la razón, y un contenido, los hechos, que se conoce mediante la experiencia. Cuando el investigador aborda áreas problemáticas poco conocidas o que se empiezan a explorar se encuentra con datos aislados, y por ello formula relaciones aisladas, sin conexión entre sí, y que no siempre están contrastadas empíricamente.

Algunas áreas educativas, como las referidas a la salud mental del educador, informática y currículum, inmersión lingüística y tratamiento de la diversidad, son relativamente novedosas. En las primeras investigaciones que se llevaron a cabo, las aportaciones no se enriquecían unas a otras ni estaban ordenadas o sistematizadas, de tal manera que desconocíamos muchas relaciones entre los fenómenos implicados y sus explicaciones. A medida que se ha desarrollado la investigación sobre los ámbitos mencionados (Esteve, 1988; García Ramos, 1988) se han descubierto y contrastado hipótesis hasta entonces aisladas. Se han establecido conexiones entre las diversas relaciones, ordenándolas coherentemente hasta formar un todo unitario que describe y explica lo mejor posible el fenómeno. En la medida que se llegue a una cohesión o encadenamiento de relaciones entre los conceptos (constructos) se irá configurando un sistema, y el conjunto que resulte de ese encadenamiento recibirá el nombre de *teoría.*

Cuando el científico elabora leyes y teorías suele recurrir a conceptos o abstracciones formadas por generalización a partir de fenómenos particulares. Así, una «agresión»

puede ser un concepto, es decir, una abstracción que identifica acciones verbales (insultos) y físicas (golpes) que tienen la característica común de estar dirigidas a despreciar o dañar a otro sujeto. Si observamos que la «agresión» suele producirse con más frecuencia en un grupo de sujetos sometidos a «frustraciones» podríamos hipotetizar la siguiente relación entre los conceptos «agresión» y «frustración»: la «frustración» tiende a producir «agresividad». Como es evidente, esta relación entre los conceptos mencionados debería ser contrastada empíricamente para poder integrarse en una teoría.

Conceptos como «agresividad», si se crean para explicar fenómenos con fines científicos, se denominan *construcciones hipotéticas* o *constructos inobservables directamente*, porque suponemos que aglutinan una serie de características que podrían ayudamos a explicar un fenómeno. Así, para explicar por qué han fracasado un grupo de alumnos podríamos recurrir a constructos como inteligencia, rendimiento, motivación y creatividad. Ninguno de estos constructos puede ser observado directamente, pero podemos observar sus manifestaciones externas, que denotan la presencia subyacente al constructo. Por ejemplo, en una conducta caracterizada por insultos y golpes, según la situación, podríamos suponer que subyace agresividad.

En vez de tratar de explicar todas las conductas específicas de carácter agresivo de los niños, el científico busca explicaciones generales. Supongamos que Carlos está jugando con un globo, Rosa lo hace explotar y Carlos insulta y golpea a Rosa. El objetivo último no es determinar si Carlos agredió a Rosa porque ésta hizo explotar el globo, sino buscar explicaciones generales que nos permitan afirmar que las privaciones tienden a generar conductas agresivas. Estas explicaciones generales constituirían lo que de forma ficticia podríamos denominar *teoría de la agresividad.* La agresividad de un niño podría explicarse por la relación mantenida con variables como el carácter, temperamento, motivación, intereses, reacción ante la frustración y características de la situación vital implicada.

Imaginemos que existiera una teoría sobre el «éxito académico». Estarían implicados constructos como inteligencia, motivación, aptitudes (verbales y numéricas) y otras variables, como horas de estudio, hábitos de estudio y conocimientos previos. El éxito académico se explicaría a partir de las relaciones que cada una de las variables tiene con el éxito académico o por las interrelaciones de ellas y el éxito académico.

Si recoge y analiza información adecuada sobre los constructos o conceptos anteriores, el científico «entenderá» y «comprenderá» mejor fenómenos como la agresividad y el éxito académico, podrá explicar por qué se producen y, hasta cierto punto, podrá predecirlos y controlarlos. Para llegar a dominar y modificar estos fenómenos, tendrá que actuar sobre variables como las características situacionales para la agresividad y sobre las horas y hábitos de estudio en el caso del éxito académico. Cabría preguntarse: ¿por qué un grupo de niños muestra agresividad? Porque se les ha obligado a diferir una gratificación muy deseada. ¿Por qué otro grupo alcanza el éxito académico en una asignatura? Porque tiene asumidos determinados hábitos de estudio ¿Por qué la gratificación diferida suele generar agresividad en los niños? Una teoría bien elaborada daría una explicación.

A partir de los ejemplos ficticios que se han propuesto puede verse que las funciones de explicación, predicción y control forman parte de las teorías. Ahora podemos ampliar más el concepto de teoría, en su sentido estricto, señalando sus elementos constitutivos (Kerlinger, 1985):

- *Conceptos o variables que describen los fenómenos. Con frecuencia se trata de constructos hipotéticos.*
- *Relaciones entre los conceptos o variables que describen los fenómenos.*
- *Explicaciones de los fenómenos descritos y de sus relaciones.*
- *Predicciones de unas variables a partir de otras.*

Veamos ejemplos reales de teorías. Algunos aspectos son tan complejos que es más adecuado hablar de distintas teorías explicativas. Así ocurre con la personalidad. Existen teorías que intentan explicar la personalidad de los sujetos y sus reacciones y manifestaciones externas, pero ninguna de ellas cumple perfectamente su cometido. Cada teoría de la personalidad enfatiza ciertas variables y presta menos atención a otras. Aún no se ha logrado una teoría de la personalidad que integre la gran multiplicidad de variables implicadas.

La teoría psicoanalítica pone de relieve que la conducta de los sujetos puede explicarse a partir de la relación que mantienen tres componentes fundamentales: Ello, Yo y Superyó. En el contexto de esta teoría, fenómenos como la ansiedad se explicarían por una relación de carácter conflictivo entre el Yo y el Superyó. La teoría de la autorrealización de Maslow propone cinco niveles de necesidades ordenadas según una jerarquía. El sujeto intenta satisfacer necesidades superiores (autorrealización, prestigio, éxito, etc.) porque ha satisfecho otras inferiores (seguridad, estabilidad, afecto, etc.). La teoría constitucional sostiene que el temperamento puede explicarse por la estructura somática, aunque las investigaciones realizadas constatan que la relación entre temperamento y estructura somática no es muy acentuada.

Características de las teorías

Básicamente, las teorías son sistemas relacionales que deben reunir características como deducibilidad, contrastabilidad y consistencia.

Sistema relaciona!. Aunque existen diversas maneras dé conceptualizar la teoría, en general se admite que está constituida por un sistema relacional de leyes que en mayor o menor grado tienden a ser generales, necesarias y constantes, estando orientadas a describir, explicar y predecir los fenómenos objeto de estudio. Además de conexionar unas relaciones con otras, la teoría trata de determinar el cómo y el por qué de las conexiones y relaciones. Es decir, da una explicación sobre determinado campo de conocimientos que ha sido explicado de manera fragmentaria por relaciones, pero que requiere una explicación integral (fig. 1.4).

Deducibilidad. Esta característica significa que es posible deducir o derivar una serie de predicciones o consecuencias de la teoría. Así, en el marco de la supuesta «teoría de la agresividad» podemos hipotetizar que la privación de una gratificación tiende a producir agresividad. Para contrastar empíricamente esta relación ha de ser posible derivar consecuencias directamente observables como la siguiente: si a un grupo de niños les impedimos el acceso a unos juguetes atrayentes que están viendo y deseando, se enfadarán e insultarán más veces que otro grupo con libre acceso a juguetes suficientes.

Esta propiedad de las teorías de poder derivar consecuencias implica que una ley puede desempeñar el papel de premisa en un razonamiento, y que se pueden derivar de ella conclusiones, por lo que es válido decir que una ley incluida en una teoría es una hipótesis (supuesto o premisa), en sentido lógico. De ahí que las teorías se denominen también *sistemas hipotético-deductivos.*

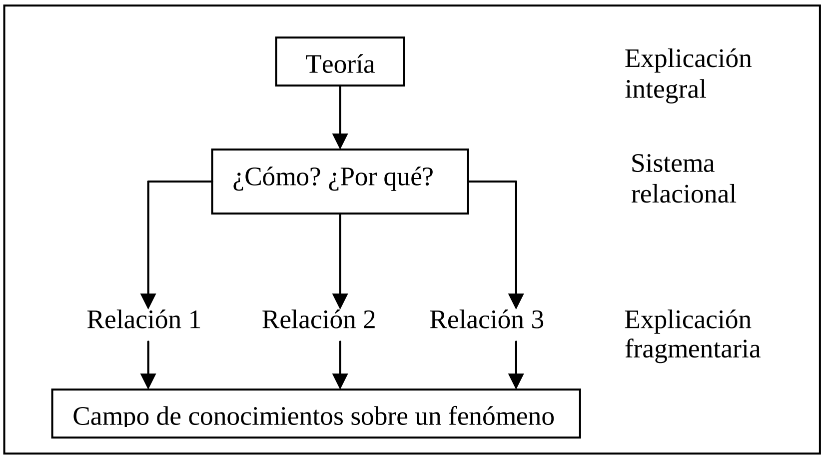


FIG. 1.4 Elementos que configuran la teoría

No es posible construir un sistema hipotético-deductivo con una hipótesis aislada. Es necesario acompañarla de proposiciones diferentes, que bien pueden ser otras hipótesis o expresiones de datos, para formar un antecedente lógico del cual se deduzcan conclusiones. Analicemos el siguiente ejemplo:

1) $p \rightarrow q$
2) p
3) $q \rightarrow r$

Antecedentes (premisa o hipótesis)

$\vdash$

4) q
5) r

Consecuentes o conclusiones

Supongamos que la premisa 1 sea la hipótesis principal; que la premisa 2 sea un dato, y que la premisa 3 sea una hipótesis subsidiaria (o que dependa de la principal). A partir de estas premisas se deducen las conclusiones 4 y 5.

De esta manera el argumento (o la expresión de un razonamiento) expresado en el ejemplo nos permite ver que la hipótesis tiene consecuencias que están realmente apoyadas en las premisas.

Si sustituimos los símbolos con proposiciones referentes a hechos (campo factual), tendremos un ejemplo como el siguiente:

1. Si se aplican incentivos positivos, entonces el rendimiento de los alumnos aumentará más que si se aplican incentivos negativos.
2. Un profesor utiliza en *sus* clases incentivos positivos.
3. Si el rendimiento aumenta más con incentivos positivos, los alumnos de dicho profesor obtendrán puntuaciones superiores a las obtenidas por alumnos sometidos a incentivos negativos.

Luego...

4. El rendimiento aumenta más con incentivos positivos que con incentivos negativos.
5. Los alumnos sometidos a incentivos positivos obtendrán puntuaciones superiores a las obtenidas por alumnos sometidos a incentivos negativos.

Contrastabilidad. Como ya se ha puesto de manifiesto al explicar la característica anterior, la contrastabilidad significa que las consecuencias derivadas de la teoría pueden compararse con la realidad para ver si contradicen los hechos observados.

Consistencia. Las teorías han de tener consistencia interna y externa. La consistencia interna se refiere al hecho de que dentro de una misma teoría no pueden existir contradicciones entre las explicaciones y predicciones que contiene. De la misma manera, al referimos a la consistencia externa queremos decir que las explicaciones y predicciones de una teoría tampoco han de estar en contradicción .con otras teorías afines.

Gracias a estas características de la teoría, las ciencias empíricas pueden cumplir sus funciones de describir, explicar y predecir los fenómenos.

Funciones de las teorías

Como ya hemos visto, el científico elabora teorías para tratar de sistematizar y aumentar el conocimiento, estableciendo conexiones lógicas entre relaciones que permiten describir, explicar y predecir la ocurrencia de los fenómenos.

La teoría posee una función explicativa y, por tanto, relacional y, a su vez, tiene una función predictiva, aunque para ello han de establecerse las condiciones en las que pueden ocurrir o no las relaciones implicadas en ella. Por su parte, la explicación y predicción intervienen conjuntamente en la planificación y aplicación racional de las acciones prácticas. Así pues, según Bunge (1984), las teorías pueden aplicarse a objetivos de conocimiento teórico o práctico. Las aplicaciones cognoscitivas de las teorías -por ejemplo, descripciones, explicaciones y predicciones- preceden a su aplicación práctica. Así, antes de tomar una decisión e implantar un cambio hay que describir cómo es el fenómeno y cuáles son sus características, hay que explicar el fenómeno tratando de averiguar por qué se produce y, por último, hay que saber cómo puede comportarse en un futuro más o menos inmediato.

Autores como Popper (1971) Y Wittgenstein (1981) consideran que las teorías son comparables a redes o mallas que nos permiten describir y captar la realidad, tratando de explicar, predecir y dominar los fenómenos. Los nudos de la malla simbolizan las relaciones entre los fenómenos, y el progreso científico consistiría en ir tejiendo una

malla cada vez más fina. La labor del científico es comparable también a la del explorador que levanta planos y mapas directamente sobre la realidad (Kaplan, 1964, y Vázquez Gómez, 1985).

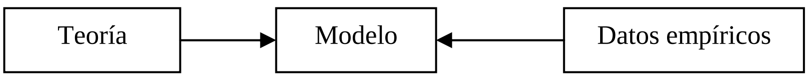
Las teorías especifican las características o variables que deben medirse antes de emprender una investigación, y aportan un lenguaje común con el que pueden enmarcarse los hallazgos de muchas investigaciones para propósitos de contrastación, comparación e investigación lógica (Wallace, 1980).

Como veremos más adelante, uno de los cometidos de la investigación científica 13s configurar y contrastar empíricamente las teorías. De ahí que una investigación sea científica en la medida que aporta información que permita generar o contrastar teorías. En consecuencia, las teorías son esenciales para la ciencia, y el proceso de la investigación científica está muy vinculado a la elaboración de teorías. Podría afirmarse que sin teoría no hay ciencia.

e) Los modelos

En el proceso de elaboración de teorías hay que tener en cuenta el papel que desempeñan los modelos. Recordemos al respecto que dos funciones básicas de las teorías consisten en explicar generalizaciones empíricas conocidas y en predecir otras que aún son desconocidas. En los procesos de teorización los modelos son elementos imprescindibles al reconocerse como representaciones orientadoras y explicativas.

A pesar de la inexistencia de un acuerdo unánime entre sus diversas conceptualizaciones, existe una tendencia general a considerar los modelos como una representación mental de un sistema real, de su estructura y de su funcionamiento. Aunque en ocasiones el término *modelo* se usa de forma equivalente a *teoría,* es común en muchas definiciones otorgarle una situación intermedia entre teoría y datos empíricos, aunque en éstos existan variaciones en matizaciones más cercanas a la teorización (Bunge, *1981b)* o a la realidad (Amau, 1978). Pero en cualquiera de los dos casos, el modelo cumple un papel de puente entre teoría y datos empíricos, convirtiéndose en un instrumento básico de investigación. Los modelos posibilitan una aproximación sistemática a los datos, componen la «base para establecer reglas de inferencia, en virtud de las cuales derivar consecuencias, empíricamente contrastables, de teorías científicas y, por fin, contribuir a explicar la teoría a partir de la cual se ha elaborado» (Tejedor, 1985b, 171).



1.4 METODOLOGÍA CIENTÍFICA

La metodología científica describe, explica y justifica el método científico. Trata de garantizar científicamente la utilización de las técnicas y estrategias implicadas en

el método científico. La ciencia se sirve de dos elementos fundamentales: los datos u observaciones y la teoría, ambos son imprescindibles para completar el ciclo de la investigación científica. Dada la complejidad y ambigüedad que entrañan estos conceptos se estudiarán más detenidamente en los párrafos que siguen.

Métodos deductivo e inductivo

Los métodos que utiliza el hombre para llegar a descubrir el conocimiento son varios. Una de las estrategias más usadas a lo largo de los tiempos ha sido el *silogismo,* cuyo razonamiento deductivo va de lo universal o general a lo particular, enlazando las premisas mayor y menor para llegar a la conclusión. Cuando el científico parte de una teoría y a través de un proceso lógico-deductivo trata de ampliarla, precisarla o contrastarla, está empleando un método deductivo.

Otro procedimiento clásico de búsqueda de conocimiento es el razonamiento inductivo, que va de lo particular a lo universal, y permite generalizar a partir de casos particulares pasando de hechos conocidos a hechos desconocidos, de conocimientos particulares a teorías o leyes generales. Cuando el científico parte de los datos y llega a la teoría emplea el método inductivo. Salvo en la inducción completa o perfecta, puede existir el riesgo de llegar a una conclusión general sobre la base de una premisa que únicamente ha estudiado cierto número de casos. Es preciso seleccionar adecuadamente los casos, evitar las generalizaciones apresuradas y reconocer que a través de esta forma se consiguen grados de probabilidad y no certeza.

Método científico

Tanto el método deductivo como el inductivo han propiciado el avance de la ciencia. No se pueden considerar como dos enfoques opuestos, sino complementarios. El modelo inductivo no puede contrastar la validez lógica de las generalizaciones empíricas a que llega, y es necesario recurrir al método deductivo. La necesidad de integrar las vías deductiva e inductiva en un único método da lugar al método hipotético-deductivo o científico.

En ocasiones, el proceso científico se inicia a través de un método inductivo. Impulsado por la necesidad de la información en tomo a una situación problemática, el científico puede partir de una serie de observaciones más o menos informales según los casos. A partir de estas observaciones exploratorias es posible realizar una recogida de datos más planificada, buscando características comunes en la información recogida. Con frecuencia se reducen los datos merced a la eliminación de información irrelevante y la búsqueda de índices. Gracias a este proceso de inducción, el científico dispone de un resumen descriptivo de los fenómenos que ha observado y de sus posibles relaciones y explicaciones.

Sin embargo, en la medida de lo posible, la ciencia busca generalizar las descripciones y explicaciones inferidas, con el fin de hacerlas extensibles a otras situaciones o hechos. Así, el científico trata de ampliar el área de conocimiento o teoría. Para ello propone un modelo, formulando hipótesis que habrán de ser contrastadas con los hechos. Este proceso parte ahora de un sistema teórico del que se desarrollan unas

premisas y conceptos que hay que hacer operativo s mediante medida o manipulación, lo que orienta la búsqueda de los datos, para contrastar empíricamente las hipótesis derivadas deductivamente de la teoría (fig. 1.5).

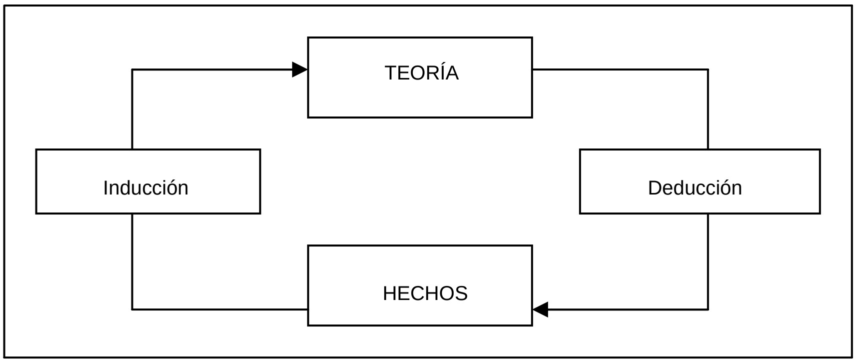


FIG. 1.5 El proceso hipotético-deductivo

El método científico de resolución de problemas es una vía o criterio para llegar a descubrir el conocimiento científico que integra la inducción y deducción. Combina, por tanto, planteamientos teóricos con la contrastación empírica de la realidad. No se trata de la simple acumulación de datos, sino de encuadrar tales hechos en una teoría e hipótesis que guíen y seleccionen el actuar empírico.

La noción de método científico que hemos expuesto sirve como referencia básica, abstracta y general que aglutina una gran variedad de posibles actuaciones. A estas especificaciones las denominamos *métodos* porque reflejan las distintas formas en las que puede o debe plasmarse en cada caso u ocasión esa noción general que llamamos *método científico.* Cuando el científico aborda problemas específicos, el método general, sin perder rigor, se hace flexible, se modifica y se adapta a la naturaleza del fenómeno objeto de estudio. Si estas adaptaciones son suficientemente importantes y generales, adquieren tal identidad que se consideran *métodos.* Como concluyen Brown y Gisselli (1969, 5), la ciencia es un método muy general que sufre diversas modificaciones, originando métodos de carácter menos general y que se utilizan en el estudio de problemas específicos.

Para Bunge (1976) el método es un modo de tratar problemas intelectuales y, consecuentemente, puede utilizarse en todos los campos de conocimiento, siendo la naturaleza del objeto en estudio la que hace aconsejables posibles métodos específicos del tema o campo de investigación correspondiente. La diversidad de las ciencias se pone de manifiesto en cuanto se llega al método general que subyace en todas ellas.

Fases del método científico

Puede considerarse que el método científico está constituido por tres *fases* o *núcleos* fundamentales (Vázquez Gómez, 1985, 163): planteamiento del problema,

construcción de un modelo que permita aproximarnos al objeto de estudio y., por último, la contrastación de dicho modelo (fig. 1.6).

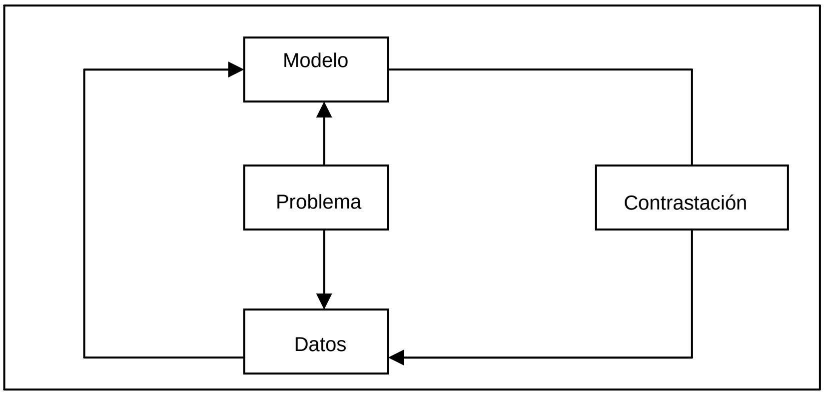


FIG. 1.6 Núcleos fundamentales del método científico

Dewey (1933, 106-118) analiza los pasos del pensamiento reflexivo y desglosa los núcleos mencionados en distintas etapas. El planteamiento del problema comprende las etapas de percepción de la dificultad y su identificación:

1) Percepción de una dificultad, obstáculo o problema para el que no se dispone de conocimientos suficientes que permitan su resolución.
2) Identificación del problema o dificultad que hay que resolver.

Identificando el problema se elabora un modelo, con una dimensión teórica (hipótesis) y otra empírica (consecuencias), es decir:

3) Planteamiento de hipótesis anticipando las respuestas a tales preguntas o la solución a las dificultades o problemas.
4) Deducción de las consecuencias de las soluciones propuestas; tales consecuencias deberían ser formuladas con precisión y en forma operativa para que sean susceptibles de observación.

Las consecuencias observadas hacen posible la contrastación del modelo propuesto.

5) Validación de las hipótesis mediante la oportuna prueba o contraste.

De modo esquemático, en el método científico pueden distinguirse unas etapas que se escalonan desde las observaciones de los datos empíricos hasta la formulación de teorías, y desde éstas hasta la contrastación de hipótesis.

1.5 INVESTIGACIÓN CIENTÍFICA

La investigación es una actividad humana realizada individualmente o en grupo, que se caracteriza por su afán persistente de conocer el mundo en que vivimos. Descubrir y conocer nuevas cosas ha sido una tarea constante en la historia de la humanidad. El pensamiento científico se ha ido gestando y perfilando históricamente por medio de un proceso que se acelera notablemente a partir del Renacimiento.

Hasta el siglo XVIII la investigación no fue objeto explícito de atención a pesar de ser tan antigua como la actividad intelectual del hombre. Desde los inicios de la filosofía griega hasta el Renacimiento, los estudios, reflexiones o especulaciones acerca de la investigación se centraron en tomo a los problemas de la lógica y del método, entendido éste como procedimiento o proceso.

El concepto de investigación se hizo eficazmente operativo a lo largo del siglo XVIII y puede considerarse la fundación de la Universidad de Berlín como el episodio significativo que inició el desarrollo creciente de la institucionalización de la investigación. Organizada por Wilhelm von Humboldt en 1809, abrió sus puertas en 1810, con la promesa de libertad de investigación (Aigrain, 1949,97). Se considera que con la fundación de dicha universidad, la ciencia empieza a ocupar el puesto de la filosofía moral y la investigación adquiere nuevos conocimientos en el campo artístico, literario o científico.

La investigación científica es una actividad intelectual organizada, disciplinada y rigurosa, que se concreta en el método científico. Para llegar al conocimiento científico, dicha actividad ha de ser sistemática, controlada, intencional, y orientarse hacia la búsqueda de nuevos saberes con los que enriquecer la ciencia. Lo que distingue su naturaleza de las otras formas de conocer es el modo de proceder y el tipo de conocimiento que persigue.

Funciones de la investigación científica

La investigación científica trata de describir, comprender, explicar y transformar la realidad. Con tal fin, uno de sus cometidos es configurar y contrastar empíricamente las teorías. Por lo que una investigación será científica en la medida que aporte información que permita generar o contrastar teorías. En consecuencia, las teorías son esenciales para la ciencia, y el progreso de la investigación científica está muy vinculado a la elaboración de teorías.

La investigación, a su vez, tiene asignadas tareas claves en el progreso de la ciencia. Para Merton (1970, 103) la investigación científica va más allá del papel pasivo de contrastar y verificar teorías; hace más que validar hipótesis. La investigación desempeña un papel activo y realiza al menos cuatro funciones fundamentales que posibilitan el desarrollo de la teoría: inicia, reforma, desvía y clarifica las teorías.

El proceso cíclico de la investigación

El proceso completo de investigación puede concebirse como un ciclo- nunca acabado (Marshall y Rossman, 1989, 22), según aparece en la figura 1.7. El esquema

refleja la importancia atribuida a la resolución de situaciones problemáticas a través de la información aportada por la investigación. Los términos escritos con mayúsculas se refieren a componentes de información (datos, generalizaciones empíricas, teorías e hipótesis, nuevos datos), que se transforman sucesivamente mediante conjuntos de métodos y técnicas (Wallace, 1980), que en el esquema aparecen escritos con minúsculas. Así, los datos se transforman en generalizaciones empíricas a través del correspondiente análisis de datos.

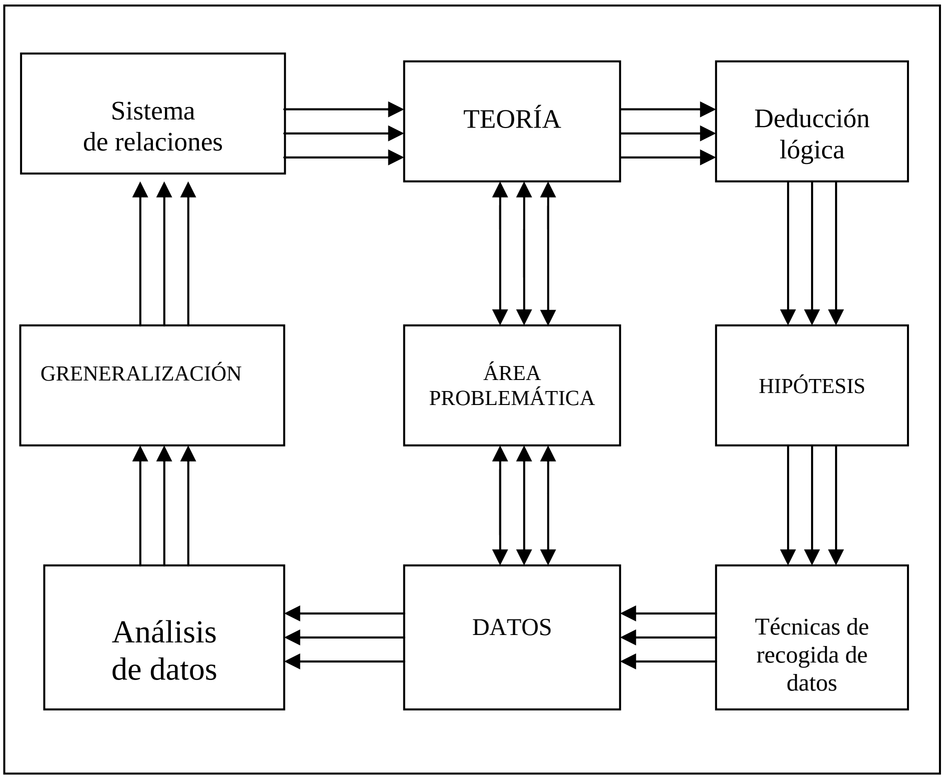


FIG. 1.7 Proceso cíclico de investigación

La explicación inductiva predomina en la mitad izquierda del esquema y la deductiva en la mitad derecha, mientras que la investigación empírica está representada por la mitad inferior y los procesos de teorización por la mitad superior (v. fig. 1.8).

En consecuencia, la investigación empírica puede utilizarse deductivamente como un método de comprobación o verificación de la teoría e inductivamente como un método de descubrimiento de generalizaciones empíricas o leyes. Ambas deben complementarse para asegurar el progreso de la ciencia.

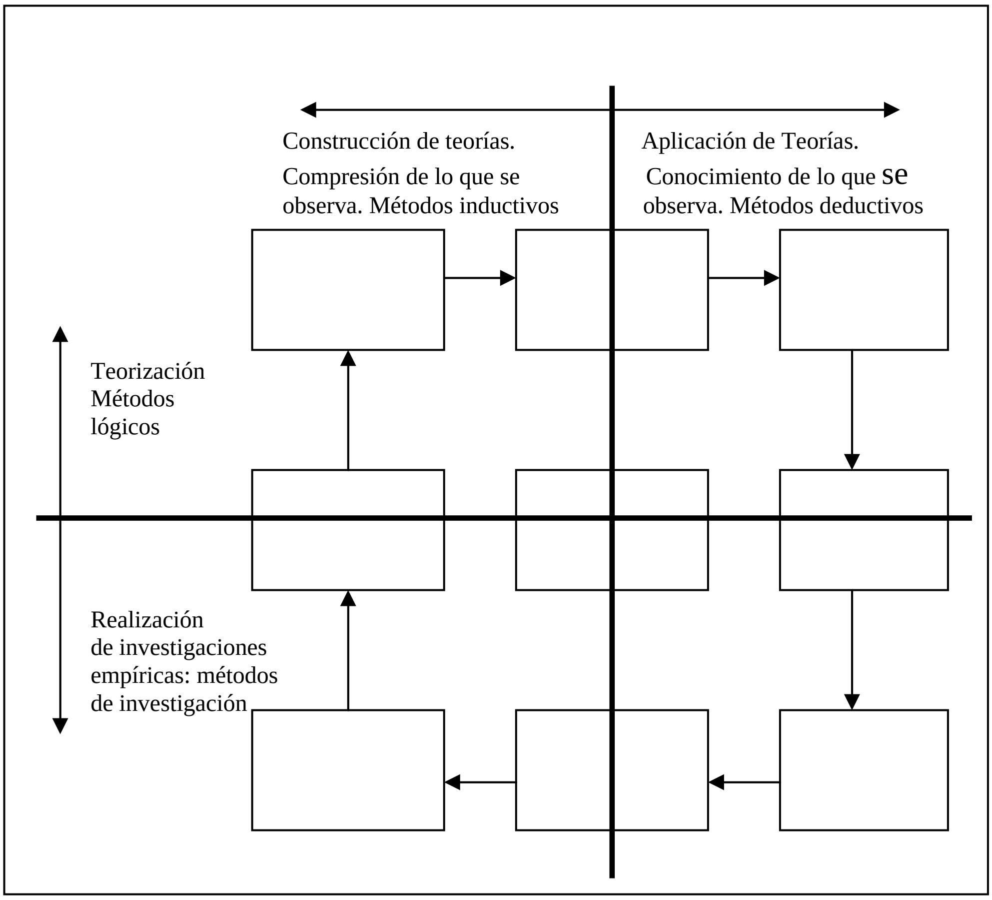


Fig. 1.8 Métodos que tienden a predominar en distintas fases del proceso cíclico de investigación. Adaptado de Wallace (1980) y Rosel (1986)

CAPÍTULO 2

Naturaleza de la investigación educativa

2.1 PERSPECTIVA HISTÓRICA

La investigación educativa es una disciplina reciente; aproximadamente tiene un siglo de historia. Sus orígenes se sitúan a finales del siglo XIX, cuando en pedagogía se adopta la metodología científica.

Antecedentes

La investigación en educación como disciplina de base empírica se denominó inicialmente *pedagogía experimental*, término equivalente al de *psicología experimental*, acuñado por Wundt en 1880.

Buyse (1949) señala tres influencias iniciales en la pedagogía experimental: el pensamiento filosófico dominante en siglo XIX, el nacimiento de la pedagogía científica y el desarrollo de la metodología experimental.

El pensamiento filosófico del siglo XIX

Las corrientes filosóficas predominantes durante el siglo XIX, como el positivismo (Comte, 1778-1857), el sociologismo (Durkheim, 185E-1917), el pragmatismo (James, 1840-1910) y el experimentalismo (Dewey, 1859-1952), son decisivas para la emancipación de las ciencias sociales y tienen importantes repercusiones en el desarrollo del método experimental. Influyen en la forma de concebir la naturaleza de los fenómenos educativos y contribuyen en buena medida a configurar el carácter científico de la pedagogía.

Nacimiento de la pedagogía científica

Otro aspecto importante que influye en la pedagogía experimental lo constituye el nacimiento de la pedagogía como ciencia. La idea de una pedagogía científica, basada

en la experimentación, se inicia con el racionalismo del siglo XVIII, la expansión de las ciencias naturales con la aportación de Darwin en el *Origen de las especies (1859)*. Obras como las de Coumot (1851), *Ensayo sobre los fundamentos de nuestros conocimientos*, y de Bain (1879), *Educación como una ciencia*, propician el nacimiento de la investigación pedagógica de carácter empírico (De Landsheere, 1985).

La publicación en 1865 de la obra de C. Bemard *Introducción al estudio de la medicina experimental* constituye un acontecimiento cumbre que inicia la aplicación del método científico-experimental en las ciencias biológicas. Los trabajos de Galton (1982) sobre biometría y procedimientos estadísticos en su laboratorio antropométrico, y las aportaciones de Burt, J. M. Cattell, Hall, Judd y Rice en el campo de la experimentación y medida de .los fenómenos psicopedagógicos hacen progresar la psicología científica, lo que favorece la aparición de una pedagogía científica (Ansoff, 1986, 19).

Las aportaciones innovadoras del pensamiento pedagógico de Rousseau (1712-1778), Pestalozzi (1746-1827) y Froebel (1782-1852) sientan las bases empíricas de la educación. Se atribuye a Herbart (1776-1841) la construcción de una ciencia de la educación fundamentada en la filosofía práctica y en la psicología, ambas con un marcado componente experimental. Posteriormente se publican obras sistemáticas que establecen los fundamentos de la pedagogía científica (Lay, 1910; Claparède, 1909; Meumann, 1914, y Buyse, 1935).

Desarrollo de la metodología experimental

La metodología experimental se desarrolla primero en ciencias afines, como la medicina y la psicología, para extenderse luego a la educación. Es el resultado de múltiples factores culturales, sociales y políticos, pudiéndose indicar como fuentes fundamentales:

a) La obra de C. Bemard (1865) expone los principios esenciales del método experimental, aclara y precisa el sentido y alcance de la metodología.
b) La consolidación de la psicología experimental con Wundt y su posterior desarrollo en los laboratorios propicia la vinculación de la pedagogía a esta forma de trabajar. Surge como pedagogía experimental al amparo de la psicología conductista, que se basa en experimentos y observaciones sistemáticas, y comienza a ser concebida como el fundamento científico preferente de los estudios de educación (T. Husén, 1988,48).
c) La aparición de las obras de Lay y Meumann. En 1905 publican conjuntamente una revisión titulada *Pedagogía experimental* haciendo hincapié en la dimensión estrictamente científica y cuantitativa del laboratorio. Meumann y Lay abordan los fenómenos educativos desde una perspectiva experimental.

Las obras de pedagogía experimental publicadas en Europa hasta la primera guerra mundial no son obras metodológicas, sino intentos de recopilar resúmenes de experimentos y trabajos científicos sobre materias escolares y la escuela. Aparte de Lay y Meumann en Alemania, el área francófona ofrece los trabajos de Claparède

(Psicología del niño y pedagogía experimental, 1905) en el Instituto J. J. Rousseau en Ginebra, defensor del experimentalismo y promotor del espíritu científico; de Van Biervliet *(Primeros elementos de pedagogía experimental,* 1911-1912, y *Principales aplicaciones de la pedagogía experimental,* 1913), de Simon *(Pedagogía experimental,* 1914).

En síntesis, el nacimiento de la pedagogía experimental tiene lugar en el contexto cultural y se relaciona con tres aspectos:

a) la preocupación por asentar la educación sobre bases empíricas;
b) la introducción del método experimental en las ciencias humanas;
c) la vinculación inicial con la psicología científica que aporta:

- un modo de trabajar: el laboratorio;
- un área de interés: los estudios psicopedagógicos;
- unos instrumentos de medida: las pruebas objetivas;
- un foco de interés: el niño (tradición paido1ógica).

Auge de la investigación cuantitativa

Aunque estas raíces científicas son imprescindibles para el desarrollo de la pedagogía experimental, según P1anchard (1960) hay que tener en cuenta también las condiciones sociales y políticas que propician un fuerte impulso de la investigación en educación, como la creación de institutos y entidades que se destinan al estudio de temas psicopedagógicos (Instituto Psicopedagógico de Leipzig, los Institutos Pedagógicos y Paidológicos de Schyten y Joteyko (Bélgica), los Institutos ,de Spearman y Burt (Inglaterra) y los de Chicago y Harvard). Progresivamente fueron apareciendo nuevas orientaciones de investigación (García Hoz, 1975):

1) la médico-pedagógica, que aborda fenómenos fisiológicos vinculados a las tareas escolares, como la memoria inmediata y la fatiga;

2) la paidológica, centrada en el estudio científico del desarrollo infantil;

3) la metodológica, interesada por la expresión cuantitativa de los fenómenos y el desarrollo de las técnicas de observación y medida.

En estas corrientes subyace un interés por lo educativo a la par que un desarrollo de las técnicas cuantitativas que se aplican ala educación. Este desarrollo metodológico afecta básicamente a la estadística ya las técnicas biométricas y psicométricas (Buyse, 1949,601).

El auge de la investigación cuantitativa puede estructurarse alrededor de importantes áreas de conocimiento como son (De Landsheere, 1988):

La estadística social

En tomo a la escuela «probabilística» francesa aparecen aportaciones como el cálculo de probabilidades, la teoría de los errores y la curva normal, debidas a científicos como Laplace y Gauss. Quetelet, a principios del siglo XIX, propone aplicar

la curva de probabilidades de Gauss a los fenómenos sociales y Galton es el primero en utilizar ampliamente la curva normal para el estudio de problemas psicológicos.

Los trabajos biométricos ingleses de Pearson, Spearman y Student desarrollan las técnicas de correlación que permiten las inferencias estadísticas. En 1904 Spearman sienta las bases del análisis factorial y Student, en 1908, diseña fórmulas para medir el error estándar de las medias y la «prueba t».

En 1903 Schuyten utiliza un diseño experimental con grupos de control, y en 1916 McCall recomienda los diseños aleatorizados. Fisher perfecciona las técnicas inferenciales y la prueba Chi-cuadrado de Pearson, añadiendo el concepto de grados de libertad; elabora el análisis de varianza y en 1935 diseña experimentos que en la década de los cuarenta se generalizarían a la educación.

En general, la estadística descriptiva paramétrica y la necesidad de estimar la significación estadística de diferencias aparece en los textos de la época (Landsheere, 1988; Thorndike, 1913; McCall, 1923; Otis, 1925; Thurstone, 1928; Yule, 1911; Brown y Thomson, 1921, en el Reino Unido; Claparede, 1909, en Suiza, y Decroly y Buyse, 1929, en Bélgica).v

Tendencias de medida

A finales del siglo XIX aparecen los primeros trabajos psicoestadísticos: J. M. Cattell (1890) desarrolla el test mental; Rice (1895) inicia el test de instrucción; Binet y Simon (1905) elaboran la «escala métrica para medir la inteligencia», y Thorndike (1910) construye la primera escala métrica (escritura), seguida por otras escalas aplicadas a distintos contenidos escolares.

En 1917 los Army tests suponen un rápido avance en técnicas de medición, de modo que alrededor de 1930 la construcción de tests estandarizado s está plenamente desarrollada. Estos instrumentos, en especial la escala de Binet, se popularizan en los países industrializados.

Encuestas administrativas

En 1817 Jullien funda la educación comparada al diseñar un cuestionario que cubre aspectos de los sistemas nacionales de educación. A finales del siglo XIX Stanley Hall y Rice perfeccionan la técnica del cuestionario en los Estados Unidos al entrevistar a profesores sobre el currículum y los métodos de enseñanza.

En 1907 Thorndike utiliza la encuesta para el estudio de las características escolares. En Europa se realizan encuestas a menor escala en Bruselas bajo los auspicios de la Escuela de Decroly y el Consejo Escocés para la Investigación en Educación sobre la capacidad intelectual.

*Desarrollo del currículum_*

El ámbito curricular es centro de interés de la investigación en educación desde sus inicios. Ya Meumann en 1900 aborda por primera vez el análisis científico de los

contenidos escolares. Con Thordike adquiere un nuevo enfoque al considerar que los contenidos de enseñanza eran objetos de la psicología escolar. Este enfoque psicologista es compatible con la filosofía pragmatista y con el intento de racionalizar el trabajo.

El movimiento progresivo, en parte inspirado en Dewey, se mantiene vinculado al anterior, pero termina rechazando su planteamiento cuantitativo experimental. Rechaza el formalismo de Thondike y enfatiza el funcionalismo en su afán de satisfacer las necesidades individuales.

En 1918, Bobbir publica una obra sobre el currículum que contribuye a destacar la definición sistemática y operativa de los objetivos educativos. Jeunehornme en 1936 escribe el *Plan d' études* belga, elaborado al conjugar las contribuciones de la investigación estrictamente empírica y de la filosofía progresiva.

Período de la crisis económica

En la década de 1930 la crisis económica recorta los fondos para la investigación. No obstante, en el ámbito de la investigación en educación tienen lugar avances y hechos de importancia.

En 1931 se crea en Gran Bretaña el Instituto de Pedagogía de la Universidad de Londres, en cuyo seno tienen lugar gran número de investigaciones. Centros como el Scottish Council for Research in Education, National Foundation for Educational Research, Northern Council for Educarion Research aportan investigaciones de reconocido prestigio en el campo del diagnóstico de las aptitudes escolares y en la predicción del rendimiento escolar. A partir de Ros años cincuenta las implicaciones sociales comienzan a hacerse patentes y se empiezan a estudiar los efectos de la reforma sobre la igualdad de oportunidades y a valorar el efecto de programas escolares. Se consolidan revistas especializadas en esta materia, como *British Journal of Educational Psychology, Educational Research* y *Research into Higher Education.*

Alemania, que ha sido un país pionero en los estudios de pedagogía científica, no tiene un desarrollo regular debido a los acontecimientos bélicos que distorsionan la vida sociopolítica del país. Antes de 1945, y en la época en que la educación comienza a considerarse disciplina académica independiente, domina el enfoque filosófico especulativo, que convive, no obstante, con el empírico. Destacan como figuras pioneras Meumann y Stern en los institutos psicológicos. Se desarrollan especialmente algunas áreas científicas como la didáctica experimental, psicología industrial y caracterología (Spranger, Kretschmer, etc.), que se abordan y estudian en centros de investigación como el Instituto Max-Planck de Berlín.

. En Francia, la pedagogía experimental está ausente durante algún tiempo, aunque deben tenerse en cuenta autores como Binet y Simon, ya mencionados, preocupados por los problemas de la medición. A partir de 1945 se produce un cambio radical; las ciencias de la educación y la experimentación alcanzan un rápido desarrollo, creándose numerosas secciones psicopedagógicas en varias universidades (París, Lyon). Como centro destacado de investigación destaca el Institut National de Recherche Pédagogique.

Bélgica es uno de los países donde más se difunden los estudios de pedagogía. En 1919 en la Universidad de Bruselas, y en 1923 en la Escuela Superior de Pedagogía de Lovaina, destaca la obra de Buyse (1935) *L' expérimentation en pédagogie,*

considerada como una de las aportaciones más relevantes de los años treinta. En la actualidad De Landsheere, en la Universidad de Lieja, continúa la tradición experimental belga En la primera mitad del siglo XX los apoyos estatales son escasos y aislados. La aportación de los estudios experimentales a la política educativa es más bien escasa, pero muy valiosa en los siguientes ámbitos:

1) técnico: diseñan procedimientos e instrumentos básicos para el estudio científico de la educación;
2) político-social: demuestran que la investigación empírica puede aportar información básica para la toma de decisiones en educación.

La segunda fase es de expansión, tanto de instituciones como de aportaciones empíricas, y se generaliza el apoyo estatal a la investigación en educación. Los enfoques que predominan son los procedimientos empíricos y cuantitativos.

Época estadounidense

La obra de Mialaret, *La nueva pedagogía científica*, publicada en 1954, es una de las últimas aparecidas en Europa que sigue manteniendo la perspectiva clásica, que basa la pedagogía experimental en el método experimental aplicado a la educación y referido fundamentalmente a las materias escolares.

A partir de los años sesenta la tradición estadounidense empieza a imponerse; la obra de Buyse se sustituye por las traducciones de Best, Travers, Van Dalen y Meyer, y más tarde por las de Kerlinger y Fox. En la Europa francófona la americanización se inicia en 1964 con la *Introducción a la investigación pedagógica,* que pública De Landsheere sin el calificativo de *experimental.*

Esta época de los sesenta se caracteriza por:

1) un clima intelectual de creación y de potenciación de alternativas;
2) una continuación de la racionalización y tecnificación del sistema escolar, lo que se traduce en una gran proliferación de tests y pruebas objetivas;
3) el apoyo de la investigación en educación por parte de la Administración;
4) un gran desarrollo de las técnicas de análisis de datos y la generalización del uso de los ordenadores.

Se potencia la organización del plan de investigación en torno a importantes instituciones como la AERA (American Educational Research Association), que se encargan de difundir los resultados de la investigación, gracias a un mayor apoyo económico.

Se genera Una extensa producción investigadora que queda recogida y sintetizada en manuales *(handbooks)* que tienen una publicación periódica, editados por la AERA (Gage, 1963; Travers, 1973; Wittrock, 1986), y obras globales como la *Encyclopedia of Educatiqnal Research,* además de numerosas publicaciones periódicas de revistas como *Joumal of Experimental Education, Joumal of Research and Development in Education, Review of Educational Research* y *Joumal of Educational Statistics.*

Posteriormente, la obra de Campbell y Stanley (1966) *Diseños experimentales y*

cuasi-experimentales en la investigación social supone un avance sustancia1 en las técnicas experimentales aplicadas a la educación. Con los diseños cuasi experimentales (Cook y Campbell, 1979) y de sujeto único (Kratochwill, 1978) se consigue una mejor adaptación a las situaciones educativas. Por otra parte, empiezan a tomar cuerpo otras orientaciones, como la etnográfica y sociológica de carácter cualitativo, que darán origen al debate metodológico.

En otros países, el avance de la investigación cuantitativa propicia una etapa de fuerte crítica de sus fundamentos sociotécnicos y la aparición de discursos metacríticos, particularmente en Francia, Alemania y el Reino Unido. La ideología estadounidense es cuestionada por otras tendencias de carácter humanista que resaltan la calidad, la persona, los derechos y la participación más que la cantidad, el grupo, las necesidades y los valores sociales.

En los años sesenta tiene lugar la crisis mundial de la educación, ante la cual aparecen distintas posiciones, como la anarquista, la democrática, la desescolarización, etc.

Alemania es el escenario de corrientes de pensamiento como la escuela de Franckfurt, de enfoque crítico, y la filosofía humanista de Husserl y Heidegger. Los paradigmas cuantitativo y cualitativo coexisten durante largo tiempo, pero con predominio del primero. En la filosofía de la ciencia se establecen debates epistemológicos (Polanyi, Popper, Kuhn, Piaget, etc.). Autores de prestigio como Cronbach y Campbell ponen en tela de juicio los planteamientos del positivismo y el predominio cuantitativo, apuntando la necesidad de métodos alternativos de investigación ..

La implantación del ordenador supone un gran avance en el análisis estadístico. Se elaboran paquetes estadísticos (BMDP y SPSS-X) que permiten no sólo realizar complejos análisis de datos, sino también crear bases de datos que posibilitan generalizar estudios bibliométricos y de metanálisis.

Tendencias durante los ochenta

Según De Landsheere (1988, 15-16), la situación actual de la investigación educativa se caracteriza por:

1) marcadas diferencias entre países en cuanto a su nivelo grado de desarrollo;
2) ampliación de cuestiones abordadas por la investigación educativa;
3) alta calidad de las investigaciones tanto en el plano teórico como en sus métodos y técnicas, lo que repercute también en la práctica educativa;
4) un status científico de la investigación educativa que alcanza un nivel de calidad comparable al de otras disciplinas;
5) reconocimiento de que ningún paradigma de investigación puede contestar adecuadamente todos los interrogantes planteados en el ámbito educativo.

Durante la década anterior el debate epistemológico había clarificado ya el alcance relativo de los enfoques cuantitativo y cualitativo.

A modo de síntesis se puede decir que la investigación educativa arraiga a nivel estatal, madura metodológicamente y amplía los contenidos temáticos. La

investigación queda ampliamente reconocida en el currículum vigente de las secciones o facultades de educación de los distintos países como fruto de su fuerte arraigo y valiosas aportaciones.

Ámbito nacional

Origen y desarrollo

Casi paralelamente a las aportaciones internacionales, en nuestro país crece el interés por la pedagogía experimental, como queda reflejado en el hecho de traducirse al castellano obras de los pioneros de esta disciplina (Lay, 1923; Meumann, 1924; Claparecte, 1927, y Buyse, 1935).

Uno de los pioneros en abordar empíricamente la educación es Galí (1928), que estudia la medida objetiva del rendimiento escolar. Otros autores vinculados a los orígenes son Mira y López, y Madariaga. A partir de 1931 se crean las secciones de pedagogía en las universidades de Madrid y Barcelona.

La guerra civil paraliza el incipiente desarrollo de esta disciplina. A partir de los años cuarenta se reanuda la actividad institucional y científica. En 1941 se crea el Instituto de Pedagogía San José de Calasanz en el seno del Consejo Superior de Investigaciones Científicas bajo la dirección de don Manuel Barbado. Desde esta institución se impulsa la investigación psicológica y pedagógica. Poco más tarde, la *Revista Española de Pedagogía* (1943) se convierte en portavoz de las aportaciones de la investigación.

En el ámbito curricular, en 1944 se incluye la pedagogía experimental y la pedagogía diferencial en los planes de estudios pedagógicos a raíz de su renovación; será García Hoz el primero en ocupar la cátedra de Pedagogía Experimental y Diferencial.

Al final de la década de los cuarenta se inician contactos con el profesor Buyse, que publica dos artículos sobre pedagogía experimental en la *Revista Española de Pedagogía* (1947-1949) y se crea la Sociedad Española de Pedagogía y la revista *Bordón* (1949), como órgano de expresión.

Junto a García Hoz (De la Orden, 1981) destacan, entre otros, Fernández Huerta, García. Yagiie, Secadas, Yela y Villarejo. Entre sus aportaciones cabe destacar:

a) tratados sobre métodos, técnicas e instrumentos de investigación que completarán los conocimientos de pedagogía experimental;
b) aportaciones de signo estadístico. Yela publica estudios sobre el análisis factorial en la revista de *Psicología General y Aplicada.*

Desde los años setenta las directrices que se observan en la universidad y en las instituciones como el CENIDE, INCIE y la red de CIE son:

a) orientación práctica, con predominio de investigaciones vinculadas a la didáctica, orientación, pedagogía diferencial y organización escolar, con la intención de cubrir demandas del sistema escolar;
b) existe claro predominio. de la metodología empírica frente a la histórica o especulativa, tanto en sus modelos como en sus métodos;

c) planificación indicativa de las investigaciones por parte de instancias superiores. Se confeccionan anualmente y con antelación los planes de investigación, en el marco de los cuales se llevarán a cabo los trabajos específicos (Escolano, 1982 y 1983).

Como características más representativas, Echeverría (1983) señala las siguientes:

a) En el ámbito universitario siguen predominando las investigaciones de orientación empírica en temáticas de signo didáctico: psicopedagogía del lenguaje, evaluación, creatividad, didáctica, orientación académica y profesional y formación del profesorado.
b) Metodológicamente cabe destacar la utilización variada de diseños de investigación y técnicas de análisis de datos que se ajustan a la temática y objetivo del estudio. La metodología descriptiva predomina sobre la experimental.
c) En 1981 se celebra el 1 Seminario de Modelos de Investigación educativa en Barcelona. A partir de estos seminarios surge la *Revista de Investigación Educativa (RIE)* en 1983, y la Asociación Interuniversitaria de Investigación Pedagógica Experimental (AIDIPE).

Durante la década de los ochenta aparecen pocas modificaciones en relación a la época anterior. En términos generales, se observa una continuidad en el interés por estudiar métodos de enseñanza y medios pedagógicos, así como temas relacionados con el rendimiento escolar. La investigación sobre el profesorado (formación, profesión docente) y sobre planificación o política educativa también suscitan en alto grado la atención de los investigadores (Ministerio de Educación y Ciencia, 1989,38).

A nivel metodológico predomina el enfoque descriptivo, apoyado especialmente en la encuesta, la observación directa y la elaboración de pruebas y/o material. Progresivamente se van utilizando métodos explicativos, predictivos y experimentales.

En síntesis, cabe destacar un creciente aumento de la investigación empírica, con un contenido muy diversificado de clara orientación práctica y con una preocupación por la dimensión metodo1ógica, ampliándose las modalidades de métodos de investigación.

España se caracteriza por una tradición y estructura investigadora poco definidas. Como señala Barto10mé (1984, 392) hay una sombra de incertidumbre sobre el futuro de la investigación educativa en España debido a:

a) falta de soporte económico suficiente y estable;
b) falta de centros educativos experimentales que vinculen teoría y práctica educativa;
c) falta de equipos de investigación que asuman programas establecidos de largo alcance.

La investigación educativa exige también una profesionalización y un marco estructural de recursos humanos y económicos estables. A pesar de ello, se está afianzando una sólida base tanto en cuanto a realizaciones científicas concretas como a apoyos institucionales.

No queremos concluir este tema sin referimos al documento *Plan de Investigación Educativa y de Formación del Profesorado* publicado por el Ministerio de Educación y Ciencia (1989) y que abre nuevas esperanzas y perspectivas sobre la investigación educativa en nuestro país. Del Plan de Investigación recogemos tres aspectos que nos parecen relevantes: los objetivos, las prioridades temáticas y las líneas de investigación.

a) OBJETIVOS DEL PLAN DE INVESTIGACIÓN EDUCATIVA '(

1. Aumento de recursos destinados al fomento de la investigación.
2. Estímulo a los equipos interdisciplinares y al empleo convergente de diversos métodos.
3. Apoyo a la labor de investigación desarrollada en los ICE y departamentos universitarios, y fomento de su relación con los CEP.
4. Inclusión de funciones relacionadas con la investigación en los CEP-
5. Estímulo a la investigación realizada P9r equipos de profesores de diferentes niveles de enseñanza y progresiva implicación de los profesores en tareas de investigación.
6. Creación de una institución que posibilite la cooperación entre las distintas administraciones.
7. Incremento de la cooperación con otros países y organismos.
8. Formación del profesorado en la tarea investigadora en el contexto de la formación permanente.
9. Estímulo a la investigación en los ámbitos de formación inicial del profesorado.
10. Aumento de recursos de difusión de la investigación.
11. Formación de especialistas en investigación y documentación educativa.
12. Desarrollo de los recursos de infraestructura documental y bibliográfica para la investigación (Ministerio de Educación y Ciencia, 1989, 70-71). ,

b) PRIORIDADES TEMÁTICAS

1. Prospectiva, planificación y política educativa en relación con los procesos de reforma del sistema.
2. Evaluación del sistema educativo y de los centros escolares.
3. Evaluación de los niveles y ciclos del sistema educativo. Definición de criterios de evaluación y validación de criterios.
4. Evaluación y determinantes del rendimiento educativo en los niveles y ciclos transformados.
5. Evaluación de programas educativos: educación especial, integración escolar, educación compensatoria y de adultos.
6. Formación y situación del profesorado.
7. Desarrollos curriculares en áreas y disciplinas específicas. Modelos curriculares.
8. Métodos, modelos y materiales didácticos.
9. Procesos de aplicación de la reforma educativa. Actitudes hacia la reforma en el medio escolar.
10. Influencia y significación de factores sociales en los procesos educativos.

11. Actitudes, valores y procesos educativos.
12. Relaciones entre el sistema educativo y el sistema productivo: educación y empleo.
13. Introducción y aplicación de nuevas tecnologías en el sistema educativo.
14. Igualdad de oportunidades en educación.
15. Modelos de orientación educativa.
16. Modelos de organización escolar.
17. Interacción educativa.
18. Participación educativa: funcionamiento de los mecanismos de participación en el sistema educativo y los centros escolares.
19. Factores cognitivos en educación.
20. Educación comparada: los sistemas educativos europeos y sus procesos de renovación.
21. Fracaso escolar.
22. Desarrollo evolutivo y educación.
23. Desarrollo y evaluación de experiencias educativas específicas.
24. Influencia educativa de los medios de comunicación social.
25. Procesos de educación no formal.
26. Pluralismo cultural y lingüístico y educación.
27. Educación y transición a la vida activa.
28. Procesos de lectura.
29. Estudios superiores: predicción del rendimiento y modelos didácticos.
30. Estudios sobre los diferentes niveles del sistema educativo (Ministerio de Educación y Ciencia, 1989,73-75)

c) LÍNEAS DE INVESTIGACIÓN

1. Prospectiva educativa.
2. Análisis comparado de los sistemas educativos en Europa.
3. Política y planificación educativa.
4. Análisis comparados de las situaciones del sistema educativo en las comunidades autónomas.
5. Situación y formación del profesorado.
6. Relaciones entre el sistema educativo y el sistema productivo.
7. Aplicación de los procesos de reforma del sistema educativo.
8. Mecanismos de participación en el sistema educativo.
9. Situación educativa de la mujer.
10. Estudios sobre diferentes niveles y ciclos del sistema educativo.
11. Formación y empleo (Ministerio de Educación y Ciencia, 1989,80).

2.2 ¿QUÉ ES INVESTIGAR EN EDUCACIÓN?

En el capítulo 1 se han tratado las nociones de conocimiento científico, ciencia, método científico e investigación científica, como fundamentos de la investigación. Ahora se aplicarán estos conceptos al ámbito de la educación, dando lugar a la diciplina denominada *investigación educativa.*

Como en el resto de las ciencias sociales, la investigación ha llegado a ser una actividad importante y necesaria en el campo educativo, dando lugar a una disciplina académica. La expresión *investigación educativa* se ha constituido, pues, en una categoría conceptual amplia en el estudio y análisis de la educación. Trata las cuestiones y problemas relativos a la naturaleza, epistemología, metodología, fines y objetivos en el marco de la búsqueda progresiva de conocimiento en el ámbito educativo.

El concepto de investigación educativa ha ido cambiando y adoptando nuevos significados a la par que han aparecido nuevos enfoques y modos de entender el hecho educativo. En la actualidad son múltiples los significados que puede asumir la expresión *investigación educativa,* si se consideran la diversidad de objetivos y finalidades que se le asignan. De ahí que virtualmente sea imposible dar una definición de *investigación* aceptada por todos o que satisfaga las diversas concepciones existentes.

En la medida que la investigación educativa acude a nuevas epistemologías o modos de conocer, surgen nuevas concepciones y significados del hecho investigador; de modo que se pueden reconocer tantas interpretaciones o concepciones del ser de la investigación educativa como tradiciones científicas existen. Se podría asumir la idea de Keeves (1988) de que existe una «unidad de investigación educativa» con diferentes enfoques mutuamente complementarios.

Una manera de acercamos a la definición es analizar cómo está concebida en las distintas perspectivas de investigación. Para autores como Best (1972), Travers (1979), Kerlinger (1985) y Ary y otros (1987), orientados hacia la corriente denominada *empírico-analítica* (positivista), investigación educativa equivale a investigación científica aplicada a la educación, y debe ceñirse a las normas del método científico en su sentido estricto.

Desde esta concepción se da importancia al carácter empírico de la investigación, apoyándose en los mismos criterios y postulados que las ciencias naturales. Sólo cuando el conocimiento se obtiene según las reglas del método científico puede ser considerado como tal, y sus leyes utilizarse en la explicación y predicción de los fenómenos. La investigación aspira a crear conocimiento teórico, cuya función radica en explicar los fenómenos educativos y eventualmente su predicción y control (Kerlinger, 1985). En educación la investigación empírico-analítica se ha preocupado mayormente de explicar las leyes de la eficacia docente.

Desde esta perspectiva, investigar en educación es *«el procedimiento más formal, sistemático e intensivo de llevar a cabo un análisis científico»* (Best, 1972, 6). Es decir, consiste en una *«actividad encaminada hacia la creación de un cuerpo organizado de conocimientos científicos sobre todo aquello que resulta de interés para los educadores»* (Travers, 1979, 19). En sentido amplio, por tanto, puede entenderse como la *«aplicación del método científico al estudio de los problemas educativos»* (Ary y otros, 1987, 20), ya sean de índole teórica o práctica.

Con el desarrollo-de nuevas ideas sobre la educación, concebida como realidad sociocultural, de naturaleza más compleja, singular y socialmente construida, han surgido nuevas conceptualizaciones de la investigación educativa, denominadas *interpretativa* y *crítica,* de corte antipositivista, y que suponen un nuevo enfoque en el estudio de la educación (cf. Goetz y LeCompte, 1984; Lincoln y Guba, 1985; Erickson, 1986; Taylor y Bogdan, 1986). La educación se concibe como acción intencionada, global y contextualizada, regida por reglas personales y sociales y no tanto por leyes científicas.

Para la concepción interpretativa, investigar es comprender la conducta humana desde los significados e intenciones de los sujetos que intervienen en el escenario educativo. Desde esta perspectiva el propósito de la investigación educativa es interpretar y comprender los fenómenos educativos más que aportar explicaciones de tipo causal.

Desde la corriente crítica o sociocrítica se destaca el compromiso explícito con la ideología y se rechaza la neutralidad del investigador. A través de la investigación aspira a transformar la sociedad en base a una concepción democrática del conocimiento y de los procesos que lo generan mediante la participación de las personas implicadas. La investigación trata de desvelar creencias, valores y supuestos que subyacen en la práctica educativa. De ahí la necesidad de plantear una relación dialéctica entre teoría y práctica mediante la reflexión crítica. De esta manera el conocimiento se genera desde la praxis y en la praxis. La investigación se concibe SQIDO un medio permanente de autorreflexión.

Según posturas más recientes, la investigación educativa, sin olvidar su función de crear conocimiento, asume el propósito de generar conocimiento útil para la acción_ educativa, ya se trate de una acción política o de un cambio en la práctica educativa (Keeves, 1988); se erige, pues, como guía de la acción educativa.

Se contempla la investigación como una *«indagación sistemática y mantenida, planificada y autocrítica, que se halla sometida a crítica pública y a las comprobaciones empíricas en donde éstas resulten adecuadas»* (Stenhouse, 1984), o como una *reflexión diagnóstica sobre la propia práctica* (Elliott, 1978). Su finalidad se centra en la búsqueda de soluciones -no de explicaciones- a los problemas educativos.

En los últimos años asistimos a la defensa de una concepción de la investigación educativa más abierta, flexible, participativa y asequible a los profesionales de la educación, comprometida con la resolución de problemas planteados desde la propia realidad educativa. En consecuencia, la preocupación por hacer de la investigación educativa una actividad estrictamente científica y rigurosa va dando paso a una mayor flexibilización, que permite tener en cuenta aspectos tales como: qué investigar, quiénes deben participar, para qué investigar y cómo debe organizarse (Escudero, 1987).

2.3 CARACTERÍSTICAS DE LA INVESTIGACIÓN EDUCATIVA

La investigación educativa posee un conjunto de características singulares. La peculiaridad de los fenómenos que estudia, la multiplicidad de los métodos que utiliza y la pluralidad de los fines y objetivos que persigue son aspectos que le confieren especificidad propia a la vez que dificultan su descripción y estudio.

A continuación reseñamos algunas características inherentes a la investigación educativa frente a la investigación en ciencias naturales:

a) *Los fenómenos educativos son más complejos.* El carácter cualitativo y complejo de la realidad educativa plantea problemas difíciles de resolver. Su estudio y conocimiento resulta más difícil que el de la realidad físico-natural debido a su mayor nivel de complejidad. Aspectos importantes de la realidad educativa como las creencias, valores o significados no son directamente observables ni susceptibles de

experimentación sin que por ello se tenga que renunciar a su estudio, como postulan los defensores del positivismo.

La realidad educativa, además de compleja, dinámica e interactiva, está dimensionada por aspectos morales, éticos y políticos que se prestan más a su estudio desde planteamientos humanístico-interpretativos. Existe, por tanto, mayor riesgo de subjetividad e imprecisión en los resultados sin que por ello tengamos que renunciar a su estudio. -

b) Los fenómenos educativos plantean mayor dificultad epistemológica. En el estudio de los fenómenos educativos, al no disponer de instrumentos precisos, no se puede alcanzar la misma exactitud y precisión que en las ciencias naturales. El carácter irrepetible de muchos fenómenos educativos dificulta su replicación. Dado que en los fenómenos educativos interactúan multiplicidad de variables, su control resulta difícil. En el ámbito educativo la conducta debe contextualizarse (Guba, 1982), lo que dificulta su generalización, ya que ésta debe estar desvinculada del contexto (Zumwalt, 1982).

c) *Su carácter pluriparadigmático.* La investigación educativa no se guía por paradigmas tan unificados e integrados como los que tienen lugar en las ciencias naturales. Dispone de mayor número de perspectivas y métodos difíciles de conciliar que le confieren un carácter pluriparadigmático y multiforme (Ministerio de Educación y Ciencia, 1989).

d) *Su carácter plurimetodológico.* La peculiaridad de los fenómenos educativos -en tanto que objeto de conocimiento- frente a los fenómenos naturales, impone ciertas restricciones. Las metodologías basadas en la experimentación y observación, consideradas por algunos investigadores como las más potentes y adecuadas para el estudio de los fenómenos educativos (Kerlinger, 1985), presentan limitaciones a la hora de su aplicación al campo educativo. Estos métodos exigen un rigor que hace difícil su aplicación en sujetos humanos. Otras posiciones defienden la necesidad de metodologías no experimentales como procedimientos más acordes con la realidad educativa. De ahí la necesidad de utilizar múltiples modelos y métodos de investigación.

e) *Su carácter multidisciplinar.* A su vez los fenómenos educativos pueden contemplarse desde diferentes disciplinas como procesos psicológicos, sociológicos o pedagógicos, lo que hace que tengan que abordarse desde una perspectiva multidisciplinar. Su estudio precisa del esfuerzo coordinado de varias disciplinas.

f) *La relación peculiar entre investigador y objeto investigado.* Otra característica es la peculiar relación que se establece entre el investigador y el objeto de estudio. El investigador forma parte del fenómeno social que investiga, la educación, y como persona que participa en él con sus valores, ideas y creencias, hace que no pueda ser totalmente independiente y neutral respecto a los fenómenos estudiados, lo que no supone renunciar en la medida de lo posible a la objetividad.

g) *Es más difícil conseguir los objetivos de la ciencia.* La variabilidad de los fenómenos educativos en el tiempo y el espacio dificulta el establecimiento de regularidades y generalizaciones, que es una de las funciones de la ciencia. Esta circunstancia obliga a adoptar posturas más prudentes que en otras ciencias.

h) *Su delimitación.* Por último, es preciso señalar que el concepto de investigación educativa no tiene un marco claro y definido para delimitar lo que puede considerarse propiamente investigación educativa. Las propuestas de innovación de métodos, modelos didácticos, pautas de interacción en el aula, etc., no son por sí mismas investigaciones educativas si no van acompañadas de procedimientos que permiten evaluar objetivamente los resultados de las innovaciones y controlar sus efectos. Su carácter difuso y relativamente impreciso obliga a mantener una actitud abierta hacia sus diferentes formas y posibilidades y a realizar un esfuerzo de clarificación (Ministerio de Educación y Ciencia, 1989,20).

2.4 PARADIGMAS DE INVESTIGACIÓN EDUCATIVA

Durante las últimas décadas, tanto en las ciencias sociales como en las ciencias de la educación, hemos asistido al surgimiento de múltiples lenguajes científicos, de pluralidad de posiciones epistemológicas y de nuevas perspectivas de investigación que se engloban bajo la denominación de *paradigmas de investigación.*

El tratamiento que reciben aquí los paradigmas es de carácter didáctico, como marco general de referencia y categoría organizadora de los principios, postulados y valores por los que se rige la diversidad de enfoques de investigación. En este sentido el concepto de paradigma nos es de gran utilidad.

Si bien el concepto de paradigma (Kuhn, 1971) admite pluralidad de significados y diferentes usos, aquí lo referimos al *conjunto de creencias y actitudes, como una visión del mundo «compartida» por un grupo de científicos que implica, específicamente, una metodología determinada* (Alvira, 1982, 34). El paradigma es un esquema teórico, o una vía de percepción y comprensión del mundo, que un grupo de científicos ha adoptado. Como se puede deducir, cada comunidad científica participa de un mismo paradigma y constituye así una comunidad intelectual cuyos miembros tienen en común un lenguaje, unos valores, unas metas, unas normas y unas creencias (Fernández Díaz, 1985, 184).

Tradicionalmente el tema de los paradigmas y su correspondiente debate se ha tratado dicotómicamente: metodología cuantitativa frente a metodología cualitativa; explicar frente a comprender; conocimiento nomotético frente a conocimiento ideográfico; investigación positivista frente a investigación humanista. Esta dicotomía deriva de las dos grandes tradiciones filosóficas predominantes en nuestra cultura: realismo e idealismo.

En el siglo XX la investigación educativa ha estado presidida por concepciones, conflictos y debates paradigmáticos; se ha movido desde posiciones dominadas por la perspectiva positivista a posiciones más pluralistas y abiertas. La era pospositivista actual (reformulación de los principios del positivismo [Guba, 1985, 15]) se caracteriza por una aceptación de la diversidad epistemológica y la pluralidad meto do lógica.

Ante el problema paradigmático se plantean diversas posiciones:

a) *incompatibilidad de paradigmas* (Smith y Heshusius, 1986);
b) *complementariedad de paradigmas* (Cook y Reichardt, 1986);
c) *unidad epistemológica* (Walker y Evers, 1988).

En la actualidad son varios los autores (Bredo y Feinberg, 1982; Koetting, 1984; Popkewitz, 1984; Soltis, 1984; Lincoln y Guba, 1985; Morin, 1985; De Miguel, 1988; entre otros) que han definido e identificado tres grandes paradigmas como marcos generales de referencia de la investigación educativa (tabla 2.1), superando la dicotomía tradicional planteada en términos de paradigma cuantitativo frente a paradigma cualitativo.

Si bien la terminología para denominar a los paradigmas es amplia, aquí utilizaremos las expresiones de paradigma *positivista, interpretativo y sociocrítico* como categorías que recogen y clarifican mejor el sentido de las perspectivas de investigación.

Paradigma positivista

También denominado *paradigma cuantitativo, empírico-analítico, racionalista, es* el paradigma dominante en algunas comunidades científicas. Tradicionalmente la investigación en educación ha seguido los postulados y principios surgidos de este paradigma. Este enfoque se vincula a las ideas positivistas y empiristas de grandes teóricos del siglo XIX y principios del XX, como Comte (1798-1857), S. Mill (1806-1873), Durkheim (1858-1917) y Popper (Viena, 1902).

El positivismo es una escuela filosófica que defiende determinados supuestos sobre la concepción del mundo y el modo de conocerlo:

a) El mundo natural tiene existencia propia, independientemente de quien lo estudia.
b) Está gobernado por leyes que permiten explicar, predecir y controlar los fenómenos del mundo natural y pueden ser descubiertas y descritas de manera objetiva y libre de valor por los investigadores con métodos adecuados.
c) El conocimiento que se obtiene se considera objetivo y factual, se basa en la experiencia y es válido para todos los tiempos y lugares, con independencia de quien lo descubre.
d) Utiliza la vía hipotético-deductiva como lógica metodológica válida para todas las ciencias.
e) Defiende la existencia de cierto grado de uniformidad y orden en la naturaleza.

En el ámbito educativo su aspiración básica es descubrir las leyes por las que se rigen los fenómenos educativos y elaborar teorías científicas que guíen la acción educativa.

Como señala Popkewitz (1988, 66), este enfoque se puede configurar a partir de cinco supuestos interrelacionados:

a) La teoría ha de ser universal, no vinculada a un contexto específico ni a las circunstancias en las que se formulan las generalizaciones.
b) Los enunciados científicos son independientes de los fines y valores de los individuos.
   La función de la ciencia se limita a descubrir las relaciones entre los hechos.
c) El mundo social existe como un sistema de variables. Éstas son elementos distintos y analíticamente separables en un sistema de interacciones.
d) La importancia de definir operativamente las variables y de que las medidas sean fiables. Los conceptos y generalizaciones sólo deben basarse en unidades de análisis que sean operativizables.

e) La importancia de la estadística como instrumento de análisis e interpretación de datos.

TABLA 2.1 Paradigmas de investigación educativa (Koetting, 1984, 296)

Dimensión	*Positivista*	*Interpretativo*	*Crítico*
Intereses	Explicar, controlar predecir	Comprender, interpretar (comprensión mutua compartida)	Emancipar, criticar e identificar el potencial para el cambio
Ontología (naturaleza de la realidad)	Dada, singular, tangible fragmentable, *convergente*	Construida, holística, *divergente,* múltiple	Contuida, holística
Relación Sujeto/objeto	Independiente, neutral, libre de valores	Interrelación, relación influida por factores subjetivos	Interrelacionados. Relación influida por el fuerte compromiso para el cambio
Propósito: generalización	Generalizaciones libres de contexto y tiempo, leyes, explicaciones (nomotéticas): - deductivas; - cuantitativas; - centradas sobre semejanzas	Hipótesis de trabajo en contexto y tiempo dado, explicaciones ideográficas, inductivas, cualitativas, centradas sobre diferencias	Lo mismo que el interpretativo
Explicación: causalidad	Causas reales, temporalmente precedentes o simultáneas	Interacción de factores	
Axiología (papel de los valores)	Libre de valores	Valores dados Influyen en la selección del teoría, método y análisis	Valores dados. Crítica de ideología

Este paradigma lleva asociado el peligro de reduccionismo al aplicarse al ámbito educativo. Si bien permite satisfacer ciertos criterios de rigor metodológico, sacrifica el estudio de otras dimensiones sustantivas del hecho educativo como realidad humana, sociocultural e incluso política e ideológica.

Por otra parte, si bien ha creado un cuerpo de conocimiento teórico como base para la práctica educativa, se cuestiona su incidencia y utilidad para mejorar la calidad de enseñanza y la práctica educativa.

Paradigma interpretativo

También denominado *paradigma cualitativo, fenomenológico, naturalista, humanista* o *etnográfico,* engloba un conjunto de corrientes humanístico-interpretativas

cuyo interés se centra en el estudio de los significados de las acciones humanas y de la vida social (Erickson, 1986).

Este enfoque tiene sus antecedentes históricos en trabajos de autores como Dilthey (1833-1911), Rickert (1863-1936), Schutz (1899-1959), Weber (1864-1920) yescuelas de pensamiento como la *fenomenología, interaccionismo simbólico, etnometodología* y *sociología cualitativa.*

Esta perspectiva pretende sustituir las nociones científicas de explicación, predicción y control del paradigma positivista por las nociones de comprensión, significado y acción. La perspectiva interpretativa penetra en el mundo personal de los sujetos (cómo interpretan las situaciones, qué significan para ellos, qué intenciones tienen). Busca la objetividad en el ámbito de los significados utilizando como criterio de evidencia el acuerdo intersubjetivo en el contexto educativo.

Desde esta concepción se cuestiona que el comportamiento de los sujetos esté gobernado por leyes generales y caracterizado por regularidades subyacentes. Los investigadores de orientación interpretativa se centran en la descripción y comprensión de 10 que es único y particular del sujeto más que en 10 generalizable; pretenden desarrollar conocimiento idiográfico y aceptan que la realidad es dinámica, múltiple y holística, a la vez que cuestionan la existencia de una realidad externa y valiosa para ser analizada.

El paradigma interpretativo se constituye como una alternativa a la visión de la perspectiva positivista. Enfatiza la comprensión e interpretación de la realidad educativa desde los significados de las personas implicadas en los contextos educativos y estudia sus creencias, intenciones, motivaciones y otras características del proceso educativo no observables directamente ni susceptibles de experimentación.

Paradigma sociocrítico

Bajo esta denominación se agrupan una familia de enfoques de investigación que surgen como respuesta a las tradiciones positivista e interpretativa y pretenden superar el reduccionismo de la primera y el conservadurismo de la segunda, admitiendo la posibilidad de una ciencia social que no sea ni puramente empírica ni sólo interpretativa (Foster, 1980).

El paradigma crítico introduce la ideología de forma explícita y la autorreflexión crítica en los procesos del conocimiento. Sus principios ideológicos tienen como finalidad la transformación de la estructura de las relaciones sociales y se apoyan en la escuela de Frackfurt (Horkheimer, Adorno), en el neomarxismo (Apple, 1982; Giroux, 1983), en la teoría crítica social de Habermas (1984) y en los trabajos de Freire (1972) y Carr y Kemmis (1983), entre otros.

Esta perspectiva tiene como objetivo el análisis de las transformaciones sociales y dar respuesta a determinados problemas generados por éstas. Algunos de sus principios son: *a)* conocer y comprender la realidad como praxis; *b)* unir teoría y práctica: conocimiento, acción y valores; c) orientar el conocimiento a emancipar y liberar al hombre, y *d)* implicar al docente a partir de la autorreflexión (Popkewitz, 1988,75).

Desde este paradigma se cuestiona la supuesta neutralidad de la ciencia, y por ende de la investigación, a la que atribuye un carácter emancipativo y transformador

de las organizaciones y procesos educativos. El grupo asume la responsabilidad de la investigación y propicia la reflexión y crítica de los intereses, interrelaciones y prácticas educativas.

En las dimensiones conceptual y metodológica existen similitudes con el paradigma interpretativo, al que añade un componente ideológico con el fin de transformar la realidad además de describirla y comprenderla.

Esta corriente tiene un impacto considerable en algunos ámbitos de la educación, como el estudio del currículum (Appel, 1982), la administración educativa (Giroux, 1983) y la formación del profesorado (Zeichner, 1983; Popkewitz, 1984), entre otros.

En la tabla 2.2 se ofrece un resumen comparativo de las características de los tres paradigmas comentados.

2.5 MODALIDADES DE INVESTIGACIÓN EDUCATIVA

En el ámbito de la investigación educativa nos encontramos con gran profusión de clasificaciones de investigación. Los criterios de clasificación son arbitrarios y no siempre mutuamente excluyentes; suelen vincularse a aspectos significativos de la investigación como finalidad, alcance temporal, profundidad y carácter de la medida. Con fines expositivos y sin ser exhaustivos, señalaremos las más comunes.

a) SEGÚN LA FINALIDAD.

Teniendo en cuenta la finalidad que persigue, la investigación se puede dividir en básica y aplicada.

Investigación básica (pura). Se define como aquella actividad orientada a la búsqueda de nuevos conocimientos y nuevos campos de investigación sin un fin práctico específico e inmediato (De la Orden, 1985). Tiene como fin crear un cuerpo de conocimiento teórico sobre los fenómenos educativos, sin preocuparse de su aplicación práctica. Se orienta a conocer y persigue la resolución de problemas amplios y de validez general (Fox, 1981, 128). En este sentido la investigación de Piaget sobre el desarrollo de la inteligencia puede considerarse básica.

Investigación aplicada. Tiene como finalidad primordial la resolución de problemas prácticos inmediatos en orden a transformar las condiciones del acto didáctico y a mejorar la calidad educativa. El propósito de realizar aportaciones al conocimiento teórico es secundario. Un estudio sobre un método de lectura para niños con dificultades perceptivas sería un ejemplo de esta modalidad.

b) SEGÚN EL ALCANCE TEMPORAL

La investigación puede referirse a un momento específico o puede extenderse a una sucesión de momentos temporales. En el primer caso se denomina *transversal* o *seccional*, y en el segundo, *longitudinal.*

Tabla 2.2 Síntesis de las características de los paradigmas de investigación

Paradigma / Dimensión	**Positivista (racionalista, cuantitativo)**	**Interpretativo (naturalista, cualitativo)**	**Sociocrítico**
Fundamentos	Positivismo lógico. Empirismo	Fenomenología. Teoría interpretativa	Teoría crítica
Naturaleza de la realidad	Objetiva, estática, única, dada, fragmentable, convergente	Dinámica, múltiple, holística, construida, divergente	Compartida, histórica, construida, dinámica, divergente
Finalidad de la investigación	Explicar, predecir, controlar los fenómenos, verificar teorías. Leyes para regular los fenómenos	Comprender e interpretar la realidad, los significados de las personas, percepciones, intenciones, acciones	Identificar potencial de cambio, emancipar sujetos. Analizar la realidad
Relación Sujeto/objeto	Independencia. Neutralidad . No se afectan. Investigador externo. Sujeto como «objeto» de investigación	Dependencia. Se afectan. Implicación investigador. Interrelación	Relación influida por el compromiso. El investigador es un sujeto más
Valores	Neutros. Investigador libre de valores. Método es garantía de objetividad	Explícitos. Influyen en la investigación	Compartidos. Ideología compartida
Teoría/ práctica	Disociadas, constituyen entidades distintas. La teoría, norma para la práctica	Relacionadas. Retroalimentación mutua	Indisociables. Relación dialéctica. La práctica es teoría en acción
Criterios de calidad	Validez, fiabilidad, objetividad	Credibilidad, confirmación, transferibilidad	Intersubjetividad, validez consensuada
Técnicas: Instrumentos Estrategias	Cuantitativos. Medición de tests, cuestionarios, observación sistemática. Experimentación	Cualitativos, descriptivos. Investigador principal instrumento. Perspectiva participantes	Estudio de casos. Técnicas dialécticas
Análisis de datos	Cuantitativo: estadística descriptiva e inferencial	Cualitativo: inducción analítica, triangulación	Intersubjetivo. Dialéctico

Investigación transversal (seccional, sincrónica). Son investigaciones que estudian un aspecto de desarrollo de los sujetos en un momento dado. Comparan diferentes grupos de edad (G_1, G_2, G_3,..., *n)* observados (O_1) en un único momento. Por ejemplo: número de palabras leídas por minuto por las alumnas de enseñanza primaria. Las muestras se estratifican por cursos.

Investigación longitudinal (diacrónica). Son investigaciones que estudian un aspecto de desarrollo de los sujetos en distintos momentos o niveles de edad (E_1, E_2… E_6), mediante observaciones repetidas (O_1, O_2, O_3,…, ***t)***. Los estudios longitudinales se_ llaman también *de panel* si se observan siempre los mismos sujetos y *de tendencia si* los sujetos son distintos. El estudio anterior realizado con los mismos sujetos en distintos momentos sería un estudio longitudinal.

c) SEGÚN LA PROFUNDIDAD U OBJETIVO

De acuerdo con este criterio, la investigación educativa puede clasificarse en exploratoria, descriptiva, explicativa y experimental.

. *Investigación exploratoria.* Tiene carácter provisional en cuanto que se realiza para- obtener un primer conocimiento de la situación donde se piensa realizar una investigación posterior. Puede ser de carácter descriptivo o explicativo, o ambos a la vez.

Investigación descriptiva. Tiene como objetivo central la descripción de los fenómenos. Se sitúa en el primer nivel del conocimiento científico. Utiliza métodos descriptivos como la observación, estudios correlacionales, de desarrollo, etc.

Investigación explicativa. Su objetivo es la explicación de los fenómenos y el estudio de sus relaciones para conocer su estructura y los aspectos que intervienen en la dinámica de aquéllos.

Investigación experimental. Estudia las .relaciones de causalidad utilizando la metodología experimental con la finalidad de control de los fenómenos. Se fundamenta en la manipulación activa y el control sistemático. Se aplica a áreas temáticas susceptibles de manipulación y medición.

d) SEGÚN EL CARÁCTER DE LA MEDIDA

Hace referencia a los dos enfoques históricos de investigación en las 'ciencias sociales: el cuantitativo y el cualitativo.

Investigación cuantitativa. Se centra fundamentalmente en los aspectos observables y susceptibles de cuantificación de los fenómenos educativos, utiliza la metodología empírico-analítica Y se sirve de pruebas estadísticas para el análisis de datos. Es la modalidad de investigación que ha predominado en educación.

Investigación cualitativa. Se orienta al estudio de los significados de las acciones humanas y de la vida social. Utiliza la metodología interpretativa (etnografía, fenomenología, interaccionismo simbólico, etc.); su interés se centra en el descubrimiento de conocimiento, el tratamiento de los datos es básicamente cualitativo.

e) SEGÚN EL MARCO EN QUE TIENE LUGAR

De laboratorio. Se realiza en situación de laboratorio, lo que conlleva la creación intencionada de las condiciones de investigación con mayor rigor y control de la situación. Los resultados de esta investigación son difícilmente generalizables a las situaciones naturales de los sujetos.

De campo o sobre el terreno. El hecho de realizarse en una situación natural permite la generalización de los resultados a situaciones afines; sin embargo, no permite el riguroso control propio de la investigación de laboratorio. Las investigaciones llevadas a cabo en el aula se consideran de este tipo.

f) SEGÚN LA CONCEPCIÓN DEL FENÓMENO EDUCATIVO

Investigación nomotética. Pretende establecer las leyes generales por las que se rigen los fenómenos educativos, orientándose hacia explicaciones generales. Utiliza la metodología empírico-analítica y se apoya básicamente en la experimentación.

Investigación idiográfica. Enfatiza lo particular e individual. Son estudios que se basan en la singularidad de los fenómenos y su objetivo fundamental no es llegar a leyes generales ni ampliar el conocimiento teórico.

g) SEGÚN LA DIMENSIÓN TEMPORAL

Investigación histórica. Estudia los fenómenos ocurridos en el pasado, reconstruyendo los acontecimientos y explicando su desarrollo, fundamentando su significado en el contexto del que ha surgido. Utiliza el método histórico y se sirve de todo tipo de documentos. La investigación histórica describe, analiza e interpreta los acontecimientos del pasado.

Investigación descriptiva. Estudia los fenómenos tal como aparecen en el presente, en él momento de realizarse el estudio. Se incluyen en esta modalidad gran variedad de estudios cuya finalidad es describir los fenómenos en el momento en que tienen lugar (estudios de desarrollo, de casos, correlacionales, etc.).

Investigación experimental. El investigador introduce cambios deliberados con el fin de observar los efectos que producen. Dado que media un tiempo entre los cambios introducidos y los efectos observados se considera orientada al futuro.

h) SEGÚN LA ORIENTACIÓN QUE ASUME

Por la orientación que la investigación asume se divide en: orientada a la comprobación, al descubrimiento y a la aplicación.

Investigación orientada a la comprobación. Es la investigación cuya orientación básica es contrastar teorías. Emplea principalmente la metodología empírico-analítica: métodos experimentales, cuasiexperimentales, ex-post-facto. Su objetivo es explicar y predecir los fenómenos. Utiliza técnicas de análisis cuantitativo y enfatiza el contexto de justificación o verificación.

Investigación orientada al descubrimiento: Es la investigación cuya orientación básica es generar o crear conocimiento desde una perspectiva inductiva. Emplea principalmente métodos interpretativos (etnografía, interaccionismo simbólico...). Su objetivo es interpretar y comprender los fenómenos. Utiliza técnicas y procedimientos de tipo cualitativo y enfatiza el contexto de descubrimiento.

Investigación orientada a la aplicación. Investigación orientada a la adquisición de conocimiento con el propósito de dar respuesta a problemas concretos. En el marco de la intervención educativa se orienta a la toma de decisiones (investigación evaluativa) y al cambio o mejora de la práctica educativa (investigación acción, investigación en el aula).

Las metodologías de investigación que se tratan en esta obra responden a la clasificación basada en el criterio de la orientación de la investigación que hemos expuesto en último lugar.

2.6 LÍMITES DE LA INVESTIGACIÓN EDUCATIVA

La realidad educativa, objeto de estudio de la investigación, se presenta, como hemos señalado ya, enormemente compleja, cambiante y diversa. Al estudiarla nos encontramos con límites u obstáculos, algunos de difícil solución. Entre los que cabe señalar:

Límites de orden ambiental

Los límites de orden ambiental se refieren a las situaciones ambientales (condiciones del ambiente, características de los sujetos, etc.).que pueden afectar a los resultados -dé-la investigación. Nada tiene que ver con las condiciones ambientales lo que en educación se entiende por «clima del aula» (ambiente social) o las múltiples circunstancias que influyen en las decisiones de las personas (necesidad de éxito, nivel de aspiración, etc.).

En investigación educativa en general, y en especial en la empírico-analítica, son múltiples las variables que intervienen y que difieren de unos contextos a otros (edad, sexo, nivel sociocultural, etc.), lo que hace que los resultados sean aplicables al contexto estudiado y hace difícil su generalización a otros contextos.

Las condiciones ambientales en educación influyen en el proceso de generalización,

poniendo límites al alcance de los resultados de investigación, de ahí que en caso de extrapolar los datos a ambientes afines se deba proceder con cautela.

Límites de orden técnico

La medida en investigación presenta dificultades de observación y cuantificación dada la naturaleza de los fenómenos educativos, de nivel más complejo que los naturales. El mundo psíquico o interior de los sujetos (significados, motivaciones, etc.) se presenta difícil de medir y observar, debido a que no es observable directamente y hay que penetrar en él a través de sus manifestaciones.

Los instrumentos y técnicas de recogida de datos disponibles en educación (tests, cuestionarios, sistemas de categorías, etc.) no alcanzan el grado de precisión y exactitud de los instrumentos utilizados en otras ciencias (rayos X, microscopio, etc.), lo que hace más difícil el conocimiento de la realidad educativa.

En los fenómenos psicológicos o internos se suelen aplicar medidas de carácter ordinal y juicios de estimación, no obstante, los supuestos implícitos en las escalas de medida son cuestionables.

A medida que vayamos consiguiendo instrumentos potentes y precisos, es decir, válidos y fiables, esta limitación disminuirá.

Límites de orden moral

La investigación con seres humanos está limitada por condiciones de tipo moral que afectan a los sujetos que participan en ella. Los límites morales se refieren a aspectos con c1ararepercusión en las personas, cuyas consecuencias sean perjudiciales (personalidad, desarrollo físico, emocional, intimidad, etc.).

La investigación, para que sea moralmente lícita, debe respetar los derechos inalienables de la persona tal como recogen las normas dictadas por asociaciones profesionales. Nada justifica el hecho de tratar a los alumnos como «objeto de investigación», en vez de sujetos de educación.

Límites derivados del objeto

La naturaleza de la realidad educativa hace difícil su conocimiento. El problema que se plantea es si la investigación educativa ha de abarcar sólo la realidad observable y cuantificable o si ha de penetrar también en la no observable (significados, intenciones, creencias, etc.).

Para el neopositivismo, todo conocimiento válido se reduce a las ciencias experimentales, y éstas se limitan a relacionar los fenómenos con datos positivos, sin la pretensión de ir más allá del mundo empírico. Las ciencias quedan reducidas a lo verificable empíricamente, y fuera de ese ámbito no tiene cabida otro conocimiento válido de los hechos.

La fiabilidad de las ciencias positivas no es suficiente ni se extiende a toda la realidad. Desde el positivismo se puede conseguir fiabilidad sólo cuando se estudian

aspectos de la realidad que son materiales, que están sometidos a leyes constantes, que pueden someterse al control propio del método experimental; muchos problemas humanos no pueden resolverse desde el positivismo, son problemas que afectan al hombre en su sentido profundo. La ciencia no es una panacea (Bunge, 1976,49) y hemos de considerarla como la vía más adecuada para resolver algunos problemas educativos o dar respuesta a algunos interrogantes; para otros, la ciencia se queda sin respuesta.

Ante el dilema de optar por el rigor y precisión neopositivista o por el riesgo de subjetividad e imprecisión que conlleva penetrar en el mundo interior de los sujetos, algunos autores se inclinan por la segunda opción.

A pesar de los límites señalados, la investigación es posible; ahora bien, nuestro propósito es advertir al investigador sobre las dificultades que condicionan la investigación. Esto exige un mayor grado de esfuerzo y búsqueda, así como capacidad de observación, reflexión y análisis. La investigación es posible como lo demuestran los' éxitos logrados y el cuerpo de conocimiento adquirido en este ámbito. No sólo es posible, sino necesaria para avanzar y conseguir nuevas cotas de conocimiento.

2.7 DEONTOLOGÍA DE LA INVESTIGACIÓN EDUCATIVA

Las ciencias humanas estudian al hombre y su conducta desde diversos enfoques. La ética se ocupa principalmente de los deberes del hombre, de la moralidad de la conducta humana, y en consecuencia de las normas éticas del quehacer científico.

Los investigadores educativos no gozan de ningún privilegio respecto a las demás personas en relación a la ética común, y por ello deben prestar atención a los principios que guían su actividad profesional; es conveniente que se establezcan obligaciones éticas por las que se debe regir la investigación. A continuación se mencionan algunas:

a) Nadie tiene derecho a entrar en la vida de otra persona para investigar sin su consentimiento y aceptación consciente y libre.

b) El investigador sólo puede hacer uso de la información obtenida para los fines previstos y conocidos por los sujetos que han participado en la investigación.

c) Al investigar con personas se deberá tener presente que éstas tienen unos derechos que se han de respetar por encima de todo. Siempre, y hoy de forma especial, se recuerda a los profesionales de la medicina, biología, periodismo, educación, etc., la conveniencia de ajustarse a las normas éticas de conducta que exige la profesión. En investigación educativa se ha procurado tener siempre presentes estas exigencias éticas.

Algunas organizaciones como la AP A (American Psychological Association) han redactado unas normas éticas para la investigación con seres humanos. El lector encontrará una síntesis de estas normas en Fox (1981,436-441), López- Barajas Y otros (1981,63-64) Y Bisquerra (1989, 18), entre otros autores.

Además de las exigencias reseñadas (libre consentimiento, información al sujeto, dar a los datos el uso previsto, observando en todo lo demás el secreto natural), las

normas éticas consideran importante evitar todo daño innecesario, físico o mental, y el deber que tiene el investigador de asegurar la calidad científica de la investigación.

En la tarea científica existen límites derivados de la naturaleza de los problemas, de los métodos y de la condición del hombre como objeto de estudio, y junto a ellos, lo que suele llamarse «integridad profesional del investigador». Un investigador podría desvirtuar su función al despreciar la exactitud en la realización, valoración e interpretación del estudio, al manipular la información obtenida o al apropiarse del trabajo de los colegas.

Según Fox (1981,436-441), al investigar deben considerarse seis condiciones. Se recoge aquí una síntesis de ellas.

1. Al sujeto se le debe informar de la finalidad de la investigación y del uso que se va a hacer de los datos que se solicitan. También tiene derecho a decidir por sí mismo si desea facilitar o no los datos para tal finalidad y uso. En caso de sujetos deficientes se debe pedir consentimiento a sus padres o tutores.

2. Los sujetos deben conocer la naturaleza del instrumento y las condiciones de la investigación lo máximo posible, antes de la recogida de datos. En ocasiones el investigador no podrá ser totalmente sincero sin afectar a la validez u objetividad del instrumento; en tal caso la información debe darse cuanto antes una vez recogidos los datos. El investigador, en este caso, no debe mentir a los sujetos y les dirá que no puede proporcionarles más detalles sobre la investigación, pues podría influir en los datos. No hay nada que justifique la mentira. Las personas que participan en una investigación comprenden, si se les explica con sinceridad, por qué no se les puede proporcionar más información sobre un proyecto o instrumento.

3. El hecho de participar en un proyecto de investigación no ha de tener ningún efecto grave sobre los sujetos o sobre su ambiente. Se debe proteger a los participantes de incomodidades, amenazas y riesgos físicos y mentales.

4. Los sujetos consultados tienen derecho a que su información sea confidencial. El investigador debe informar a los participantes. Deberá cumplir rigurosamente las condiciones y acuerdos negociados con los sujetos y las promesas que les haga.

5. El instrumento y la situación de investigación no podrán exigir una conducta que en condiciones normales no se espera obtener de los sujetos o se les permite realizar. Esto significa que el investigador debe actuar con unas normas éticas profesionales.

6. La situación de investigación debe permitir que los sujetos obtengan el máximo provecho de su participación. La vía de la comunicación entre investigador y sujetos investigados ha de ser bidireccional. Si éstos lo desean, pueden exigir que se les proporcione un resumen de los datos o el informe de investigación.

CAPÍTULO 3

Proceso general de la investigación

Introducción

Las acciones o actividad del investigador se configuran en tomo a tres grandes etapas (fig. 3.1). La primera consiste en la elaboración del proyecto de investigación, etapa que se corresponde con la fase de planificación del proceso general de investigación; la segunda etapa la constituye la recogida y análisis de datos, correspondiente a la fase de la realización de la investigación; y, por último, la tercera se refiere a la comunicación de los resultados a través del informe.

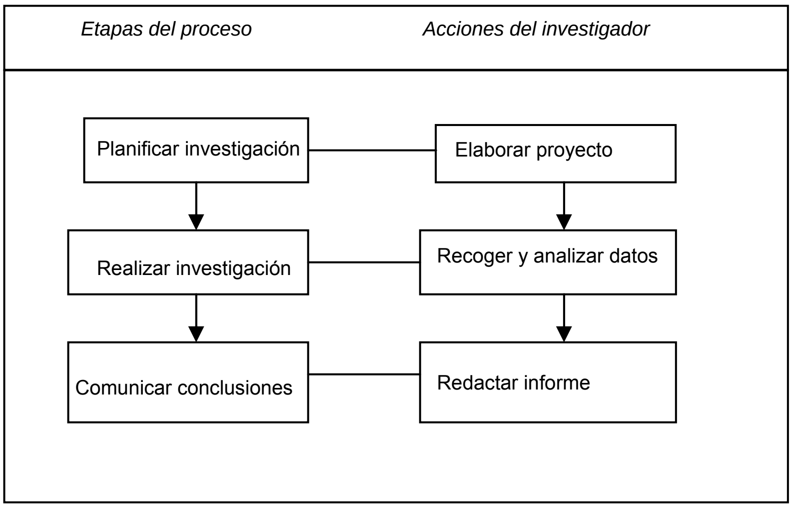


Fig. 3.1 Etapas del proceso general de investigación y acciones del investigador

Este capítulo describe el proceso general de investigación con el fin de ofrecer directrices que orienten su planificación y realización y sirvan de guía para elaborar proyectos de investigación. También se exponen con mayor amplitud las fases de: planteamiento del problema, revisión de la literatura, hipótesis y variables, técnicas de muestreo, así como una primera exposición de las diversas metodologías de investigación entre las que podrá optar el investigador para dar, respuesta adecuada al problema planteado.

3.1 PERSPECTIVA GENERAL DEL PROCESO

Antes de pasar a exponer cada una de las etapas del proceso de investigación es pertinente tener de éste una perspectiva general, lo que nos permitirá comprender mejor su estructura y funciones. El objetivo de este apartado es ofrecer una visión global del proceso de investigación como actividad unitaria y coherente que sirva de base para el estudio posterior de cada una sus etapas o momentos.

El modelo de investigación presentado (fig. 3.2) sirve para ilustrar la perspectiva general del proceso de investigación.

A continuación se realiza un breve comentario de las etapas más importantes del proceso de investigación que más tarde se desarrollarán detenidamente.

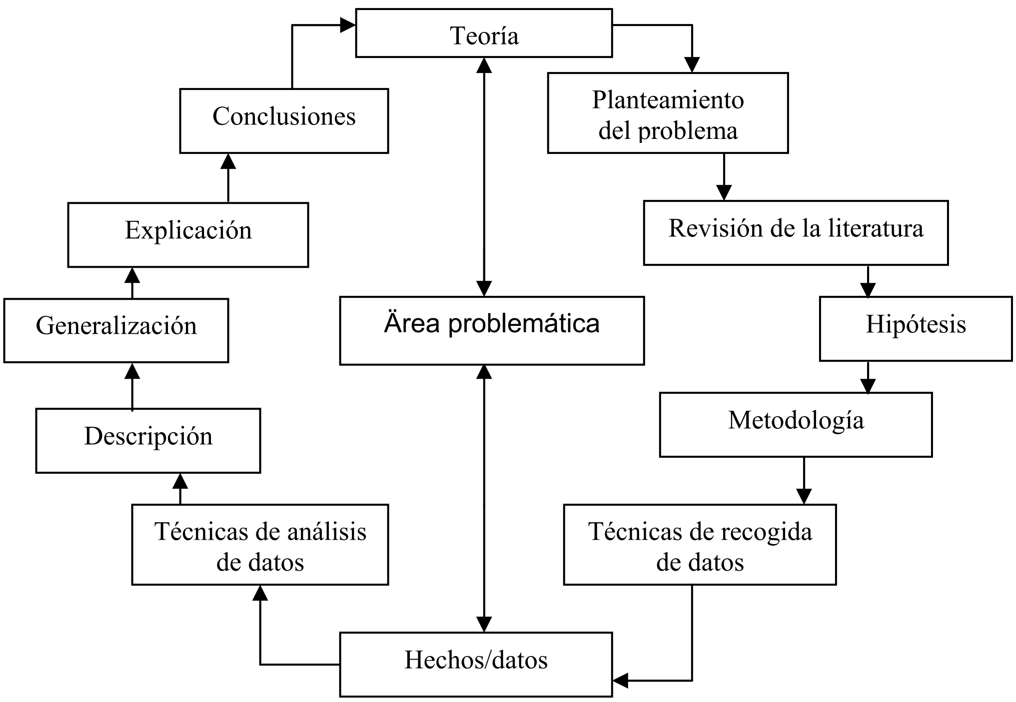


FIG.3.2 Perspectiva general del proceso de investigación

Planteamiento del problema

El proceso se inicia en tomo a un área problemática de la que se extrae el problema de investigación; éste surge de un contexto teórico o práctico. Esta etapa puede sintetizarse en la pregunta: ¿de qué se trata?

En esta etapa se consideran aspectos referidos a la naturaleza del problema, como su identificación, valoración, formulación y tipos.

La elección del problema es de índole personal; depende de los intereses y objetivos del investigador: comprobar teorías, descubrir o generar conocimiento o mejorar y optimizar la práctica educativa.

Revisión de la bibliografía

Esta etapa del proceso consiste en hacer una revisión sistemática del «estado de la cuestión» sobre el tema de investigación. El investigador consulta y revisa las fuentes, busca, recopila y organiza la documentación sobre el tema. Esta etapa es fundamental para la investigación.

Hipótesis y variables

Una vez formulado el problema y revisada la bibliografía sobre él, se procede a formular la hipótesis como una respuesta conjetural al problema. Esta etapa requiere precisar y concretar el problema en una respuesta y exige las condiciones de fundamentación y posibilidad de contrastación.

La formulación de la hipótesis de un modo claro y preciso es básico en los trabajos de investigación. Las hipótesis pueden surgir de la teoría (vía deductiva) o de la *v* práctica (vía inductiva), denominándose en este caso *microhipótesis.* Debe enunciarse a un nivel de concreción que permita medir u observar las variables.

Cuando la investigación se orienta a comprobar teorías, las hipótesis desempeñan un papel de avance o solución del problema; la orientada al descubrimiento requiere un proceso inductivo de búsqueda de información, pudiendo llegar a formularse microhipótesis. La hipótesis sirve de nexo entre la teoría y las observaciones o datos y se constituye en el momento clave de la actividad científica.

La hipótesis relaciona dos o más variables. Éstas deben expresarse en lenguaje claro y preciso, en términos empíricos, es decir, que se puedan observar o medir. El paso de un nivel conceptual a un nivel observacional se denomina *operativización de las variables.*

Metodología

Otro aspecto del proceso de investigación lo constituye la metodología, que es el plan o esquema de trabajo del investigador. Recoge los procedimientos que se seguirán para contrastar la hipótesis e incluye consideraciones respecto al método apropiado y el tipo de diseño, la selección de los sujetos.

Método de investigación

El investigador deberá elegir un método adecuado para los objetivos de la investigación. El problema puede tener una orientación de comprobación, de descubrimiento o de aplicación para la práctica. Puede ir desde la experimentación a la descripción o interpretación de los hechos. La naturaleza del área problemática y los objetivos de la investigación determinarán el método que hay que utilizar.

La misma naturaleza de los fenómenos educativos admite y precisa de una gran variedad de métodos de investigación, a los que van asociados multiplicidad de diseños y estrategias.

Diseño

El método se traduce en un diseño que refleja el plan o esquema de trabajo del investigador. El diseño describe con detalle qué se debe hacer y cómo realizarlo, plasma las actividades, incluye los grupos de sujetos, las variables implicadas. Tiene gran valor como clarificador y especificador de las ideas y tareas que hay que realizar.

Selección de la muestra

El investigador define la población adecuada objeto de estudio y extrae una muestra representativa de individuos utilizando la técnica de muestreo apropiada que garantice la representatividad de los individuos elegidos.

Técnicas de recogida de datos

Son medios que el investigador utiliza para facilitar el tratamiento de los datos propio de cada investigación. En investigación educativa disponemos de gran variedad de técnicas (tests, cuestionarios, escalas, sistemas de observación, etc.) elaborados para cubrir las necesidades de la investigación. Cada uno posee inconvenientes y ventajas y tiene diferentes usos. A la hora de elegir un instrumento, el investigador debe tener en cuenta su validez y fiabilidad.

Técnicas de análisis de datos

El propósito del análisis de datos consiste en organizar y tratar la información para que se pueda describir, analizar e interpretar. La naturaleza de los datos requiere un análisis cuantitativo, cualitativo o ambos.

Conclusiones

Esta etapa del proceso de investigación recoge los resultados del estudio. Se incluyen aspectos como la aceptación o rechazo de las hipótesis, el poder de

generalización de los datos, la metodología utilizada, coincidencias o desacuerdos con otras investigaciones, implicaciones para la práctica y sugerencias para posteriores investigaciones.

Expuesta la perspectiva general del proceso, a continuación se desarrollan cada una de las etapas mencionadas.

3.2 PLANTEAMIENTO DEL PROBLEMA

El planteamiento del problema supone el inicio del quehacer científico, aunque no tiene por qué ser exclusivamente la única manera de poner en marcha el proceso de investigación. Ésta surge cuando tenemos que dar respuesta a una pregunta o dificultad que consideramos problemática.

El problema de investigación adquiere distintos sentidos según los fines que se plantee el investigador. Puede considerarse como una *dificultad* u *obstáculo* que desconcierta a los investigadores (Dewey, 1933), como una pregunta que el investigador se plantea de cara a la comprensión y explicación de un fenómeno educativo (Bartolomé, 1983) o incluso como una situación educativa que el profesor o maestro considera como susceptible de cambio o mejora (Elliott, 1983). En sentido amplio, un problema es *cualquier situación sin una solución satisfactoria.*

El estudio del planteamiento del problema se aborda mediante la descripción de algunos de los aspectos que más lo caracterizan, como son su *identificación, valoración, formulación, tipos* y *etapas.*

Identificación del problema

¿Cómo se llega a los problemas? ¿Cómo extraerlos de una situación o área problemática? La respuesta no es fácil; en ocasiones hasta puede que resulte la tarea más compleja del proceso de investigación. No disponemos de reglas o procedimientos que nos lleven directamente a la identificación de los problemas. La realidad educativa, de por sí compleja y problemática, no se manifiesta por sí misma, a no ser que nos acerquemos a ella, en tanto que expertos o investigadores. Los problemas se identifican cuando los extraemos de esa realidad compleja en la que se hallan inmersos y conseguimos aislarlos, lo que nos permite su descripción, conocer sus características, los antecedentes que los explican, cómo están relacionados, cuál es su contexto, etc.

La elección del área problemática viene determinada por el conocimiento, experiencia y condiciones que rodean al investigador. Para un profesor el interés puede centrarse en el estudio de los métodos de lectura. Lo que puede llevarle a preguntarse por qué unos alumnos poseen mayor velocidad lectora, cómo aumentar la eficacia lectora o qué método de lectura es más eficaz para optimizar el proceso lector.

Los problemas de investigación pueden surgir de diferentes fuentes. Bartolomé (1983,41) las resume en tres: *a)* la experiencia educativa, «la fuente más rica y fecunda de situaciones problemáticas»; *b)* el campo teórico: estudio crítico de las ciencias de la educación, y *c)* la investigación pedagógica ya realizada.

Las aportaciones de Travers (1979), Pérez Juste y otros (1981, 96) señalan que un problema puede surgir como: 1) fruto de la observación; 2) resultado del contacto con

profesores o personas expertas; 3) consecuencia de lecturas de publicaciones, estudiando cómo se podrían aplicar a la solución de problemas las técnicas e ideas examinadas en ellas; 4) conclusión de las charlas o discusiones entre especialistas; 5) examen del apartado «discusiones» o «sugerencias para futuras investigaciones» que todo informe de investigación debe tener.

La experiencia docente, el interés por mejorar o cambiar la propia práctica educativa o el contraste entre lo que creemos que debe ser la educación y lo que realmente es, son fuentes donde los prácticos encuentran preguntas o dificultades a las que dar respuesta.

Valoración del problema

Valorar un problema de investigación consiste en considerar si reúne una serie de condiciones que se estiman en sí mismas necesarias.

Para los problemas que requieren un planteamiento orientado al contexto de comprobación, autores como Ary y otros (1987), Bartolomé (1982, 42-43), Van Dalen y Meyer (1983) señalan que el problema debe ser:

Real. Debe partirse de la existencia de un problema percibido o sentido.

Factible. Que reúna las condiciones para ser estudiado. Se considerará su dificultad, recursos disponibles, acceso a la información, financiación; es decir, que esté al alcance del investigador.

Relevante. El investigador debería reflexionar sobre aspectos como: el problema ¿tiene relevancia práctica?, ¿me interesa?, ¿es importante?, ¿es actual?, ¿qué soluciones aporta?

Resoluble. Un problema es resoluble si: *a)* puede formularse una hipótesis como tentativa de solución; *b)* es posible comprobar dicha hipótesis determinando un grado de probabilidad (McGuigan, 1977,37).

Generador de conocimiento. El investigador debe reflexionar si la resolución del problema contribuirá a crear conocimiento pedagógico o cubrirá alguna laguna en el conocimiento actual.

Generador de nuevos problemas. La solución del problema debe conducir a nuevos problemas e investigaciones. La respuesta a un interrogante, de ordinario, plantea otros.

Sin embargo, desde otros enfoques de investigación, como la investigación acción, a la hora de valorar el problema de investigación hacen más hincapié en criterios como: que sea práctico y real, que tenga utilidad para los profesores o alumnos, que responda a las necesidades de la práctica educativa, que sea aplicable a corto plazo o que sirva para cambiar o mejorar la práctica.

En situaciones de formación o iniciación en la investigación, los criterios deben matizarse. Se contemplarán los problemas desde un punto de vista didáctico y funcional para que el proceso de aprendizaje del hecho investigador tenga lugar. Se elegirán temas fáciles y bien conocidos.

Formulación del problema

Una vez identificado y valorado el problema se procede a su formulación, que, en especial desde una perspectiva empírico-analítica, fundamentalmente consiste en reducido a sus aspectos y relaciones esenciales.

Para Ary y otros (1987) la formulación del problema debe reunir dos condiciones: *a)* especificar lo que ha de determinarse o resol verse, y *b)* restringir el campo de estudio en un interrogante concreto. Kerlinger señala (1985) tres condiciones: *a)* debe expresar la relación entre dos o más variables; *b)* debe enunciarse en forma clara y unívoca, de modo que la solución sólo admita respuestas precisas y, a ser posible, en forma de pregunta, y e) debe ser susceptible de verificación empírica. No debe plantear juicios de valor sobre lo que es mejor o peor, sobre cómo debería ser idealmente la realidad, sino sobre cómo es realmente.

Todos los autores coinciden en destacar la necesidad de que el problema sea formulado con precisión, en una o varias preguntas concretas donde se relacionen las variables implicadas, de forma que se constituya en guía para la formulación de la hipótesis (Kerlinger, 1985).

A continuación presentamos algunos ejemplos de formulación de problemas de investigación:

1. ¿Influye la visualización de un documental fílmico sobre el tema de las centrales nucleares en España en la actitud de los universitarios con respecto a esta cuestión?
2. ¿Existe relación entre las horas pasadas delante de la televisión y la conducta agresiva de tipo verbal en adultos de 25-35 años residentes en localidades mayores de 100000 habitantes?
3. Un 35 % de alumnos presentados a las pruebas de acceso a la Universidad X (curso 88-89) han puesto de manifiesto deficiencias ortográficas en sus ejercicios escritos. ¿Qué condiciones concurren en estos alumnos que, después de doce años de escolaridad, usan incorrectamente el vocabulario usual de la lengua castellana?
4. Al iniciar su ejercicio profesional, en el proceso de transición de la formación inicial a la situación real de trabajo, los profesores noveles encuentran dificultades y problemas. ¿Cuáles son los problemas y dificultades personales que encuentran en el ejercicio de la profesión los profesores de enseñanza media según los años de experiencia?
5. ¿Los alumnos que aprenden a leer por un método fónico o silábico desarrollan iguales hábitos o estrategias intelectivos que los que aprenden por un método global-natural?

Tipos de problemas

Los problemas de investigación, como hemos señalado más arriba, pueden surgir de diferentes fuentes: de la práctica educativa, de la observación de una situación que necesita explicación o de alguna teoría. El procedimiento de búsqueda sigue vías diferentes según sea el origen del problema: la vía inductiva, si el problema surge de la realidad; o la vía deductiva, si surge de la teoría.

El procedimiento inductivo constituye una fuente particularmente útil para los prácticos. En estos casos los prob1emas surgen de la experiencia o práctica _educativa, tienen como punto de partida la reflexión del investigador sobre la realidad educativa. A este tipo de problema se le denomina *abierto* y su solución se orienta a generar conocimiento práctico.

El procedimiento deductivo constituye una fuente útil para los teóricos. Los problemas surgen de las teorías o de problemas ya planteados anteriormente. A este tipo de problema se le denomina *cerrado* y su solución se encamina a comprobar teorías (Hopkins, 1989).

El diagrama que se adjunta representa la relación entre tipo de problema, hipótesis y orientación.

Tipo	*Hipótesis*	*Orientación*
Cerrado	deductiva	comprobar
abierto	inductiva	generar

Etapas del planteamiento del problema

Cada investigador suele seguir determinados procedimientos para formular el problema de investigación. El planteamiento del problema conlleva las siguientes etapas:

1. Elección de un área problemática.
2. Identificación y delimitación del problema.
3. Valoración del problema.
4. Formulación del problema en forma de pregunta.

Gráficamente el proceso del planteamiento del problema se puede representar tal como aparece en la figura 3.3.

3.3 REVISIÓN DE LA BIBLIOGRAFÍA

Este apartado tiene por objeto presentar y describir cómo se realiza la revisión de la bibliografía sobre un tema de investigación, teniendo en cuenta que cada tema tiene su propia historia. Realizar una revisión consiste en consultar y revisar sistemáticamente las fuentes de documentación. Resulta necesario conocer aquello que se ha dicho, hecho o descubierto sobre el problema de investigación. Este apartado es de interés, entre otras razones, porque el problema podría estar resuelto ya de modo satisfactorio y no tendríamos más que recoger sus aportaciones sobre el tema. El estudio previo del «estado de la cuestión» es imprescindible.

Saber consultar la bibliografía sobre un tema o cuestión, citar correctamente,

mostrar rigor terminológico, localizar, recopilar y organizar adecuadamente la documentación sobre el problema, tiene un especial interés para el que se inicia en la investigación y para el mismo proceso de investigación.

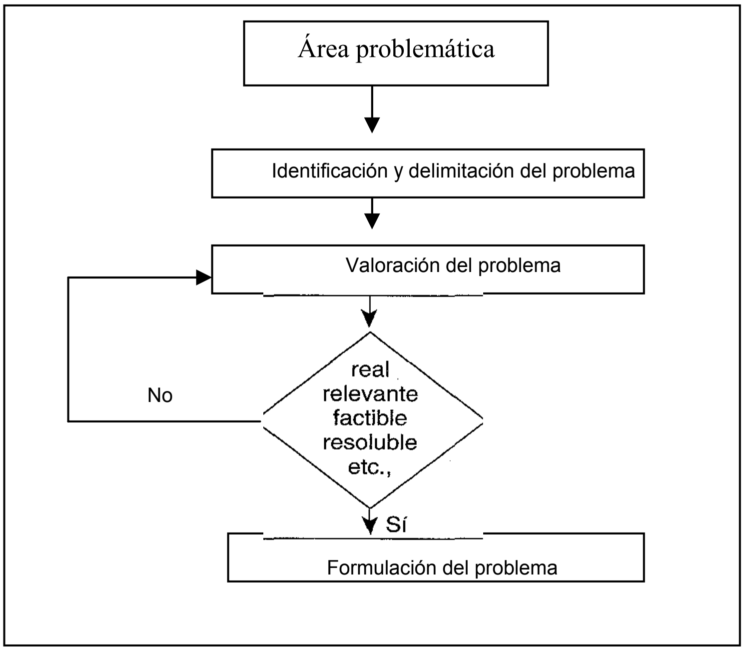


FIG. 3.3 Etapas del planteamiento del problema

Marco teórico

Una vez planteado el problema, el investigador debe elaborar el marco teórico, para lo que puede acudir a diversas fuentes, de conocimiento popular, de divulgación o científicas. La revisión de fuentes le permitirá conocer qué se sabe o se ha investigado sobre el tema.

El marco teórico expone y analiza las teorías o grupos de teorías que sirven de fundamento para explicar los antecedentes e interpretar los resultados de la investigación (Mtinch y Angeles, 1988,69). Estos autores proponen el proceso de la figura 3.4 para la elaboración del marco teórico.

Mediante el conocimiento de los aspectos teóricos referentes al problema, el investigador conseguirá una marco de referencia conceptual, indicaciones y sugerencias para el enfoque, el método y la instrumentación; la posibilidad de determinar las variables, formular la hipótesis e interpretar adecuadamente los resultados.

La construcción del marco teórico es un trabajo que se inicia al comienzo de la investigación y que se perfecciona durante todo el proceso. Conviene resaltar que establecer el marco teórico implica recoger información sobre el tema y utilizada para sentar las bases teóricas de la investigación.

En el marco conceptual se deberá contestar a ciertas preguntas como: ¿qué investigaciones se han realizado sobre el problema?, ¿quién las ha realizado?, ¿cuándo?, ¿dónde?, ¿cómo? El marco teórico debe concretarse y referirse específicamente al problema. En su elaboración, el investigador, además de presentar la información recopilada, debe presentar sus conclusiones, críticas y aportaciones acerca del problema, sugerencias sobre futuras áreas de investigación, hecho que enriquecerá el material obtenido.

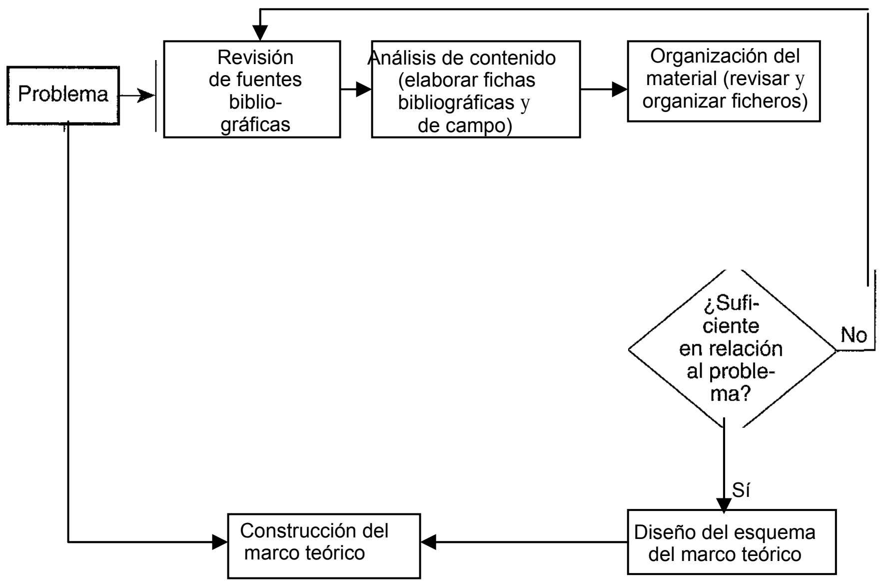


FIG. 3.4 Elaboración del marco teórico (Múrch y Ángeles, 1988, 69)

El *status quaestionis* sobre el tema debe ofrecer la recopilación ordenada de las fuentes documentales y las lagunas que existen sobre el problema en estudio. Mediante citas, notas y referencias debe darse crédito a los autores de los que se obtenga la información.

El estado de la cuestión posee tal importancia e independencia respecto al resto del trabajo de investigación, que permite su publicación aparte, aunque de ordinario aparezca en la introducción.

Revisión bibliográfica

El enorme flujo de información existente dificulta el acceso a aquel material que es relevante para el investigador. Pasamos de un sistema artes anal de fichas a un sistema informatizado, lo que exige al investigador una formación en este campo de la informática que le permita acceder a las fuentes de información. La consulta de la información se constituye en una etapa desafiante a la vez que imprescindible para la

investigación. A continuación presentamos un resumen del listado de las principales referencias bibliográficas siguiendo el esquema de Bisquerra (1989,21-29).

Fuentes bibliográficas

- Índices de las revistas especializadas en el tema.
- Diccionarios, enciclopedias, manuales, monografías, obras especializadas sobre el tema. *Enciclopedia Internacional de la Educación* (1990) o *Handbook oj Research on Teaching (1986).*
- Índices bibliográficos como: BIBE *(Boletín Internacional de Bibliografía sobre Educación),* ERIC (Educational Research Information Centers), RIE (Resources in Education), CIJE (Current Index to Journals in Education), Current contents, Education Index, CIDE (Centro Nacional de Investigación y Documentación), etc.
- Abstracts. Un abstract es una revista que recoge resúmenes sobre libros, revistas, artículos, etc.: *Educational Abstracts, Dissertation Abstracts International* (ofrece el servicio DA TRIX II de la University Microfilms, resúmenes de tesis doctorales), *Psychological Abstracts,* etc.
- Tesis doctorales, tesinas, ponencias y comunicaciones a congresos y seminarios, investigaciones inéditas.
- Centros de teledocumentación (informatizados) como: ISOC (Instituto de Información y Documentación en Ciencias Sociales), CSIC (Consejo Superior de Investigaciones Científicas), CIDC (Consorci d'Informació i Documentació de Catalunya).
- Bases de datos. La red INCA (Red de Información Científica Automatizada) permite acceder a las principales bases de datos nacionales e internacionales. Bases de datos conocidas son: ERIC (Educational Resources Information Center), RIE (Resources in Education), CIJE (Current Index to Journals in Education) EUDISED (European Documentation and Information System for Education), REDINET (Red Estatal de Bases de Datos sobre Investigaciones Educativas).
- Los thesaurus son vocabularios de palabras-clave que posibilitan el acceso a las fuentes de documentación. Se presentan en forma de una lista ordenada de descriptores (palabras-clave). Son fundamentales para la búsqueda de referencias bibliográficas.

Referencias bibliográficas

Los datos técnicos necesarios para la perfecta identificación de una cita se denominan *referencia bibliográfica.* El modo de referenciar las citas es arbitrario. Cualquiera que sea el método adoptado, su presentación debe ser uniforme y coherente en una misma obra o en una misma lista.

No es suficiente con conocer las fuentes documentales; se necesita también saber referenciar y hacer las citas correctamente. Se denomina *cita* al texto o idea de un autor que se transcribe y que aparece en el cuerpo de un escrito o en nota de pie de página.

Aquí exponemos el procedimiento recomendado en Bisquerra (1989, 23-29), siguiendo las instrucciones de la AP A (American Psychological Association, 1983).

1. LIBROS

a) Debe aparecer: apellido del autor (coma), inicial/es del nombre (punto), fecha entre paréntesis (punto), título subrayado o en cursiva (punto), lugar de edición (dos puntos), editorial. Por ejemplo:

Anguera, M. T. (1983). *Manual de prácticas de observación.* México: Trillas.

b) Cuando el lugar de edición no es una capital conocida, es apropiado citar el estado. Por ejemplo:

Agar, M. H. (1986). *Speaking of Etnography.* Beverly Hills, California: Sage.

c) El estado se puede abreviar con las dos primeras letras. Por ejemplo:

Kirk, R. E. (1982). *Experimental design: Proceduresfor the Behavioral Sciences.* Belmont, Ca.: Brooks/Cole Publishing Co.

d) Si hay más de un autor deben indicarse todos, separados por comas, excepto el último, que va precedio de la conjunción *y.* Por ejemplo:

Goetz, J. P., Y LeCompte M. D. (1988). *Etnografía y diseño cualitativo en investigación educativa.* Madrid: Morata.

e) Si a lo largo del texto se cita una referencia de más de tres autores se puede mencionar el primero seguido de la expresión *et al. (y otros).* Por ejemplo:

Alvira et al. (1979) «Bartolomé ET AL. (1982)>>.
En la bibliografía deben aparecer todos los autores. Ejemplo: Bartolomé, M., B. Echeverría,1. Mateo y Rodríguez, S. (coord.) (1982). *Modelos de investigación educativa.* Barcelona: ICE de la Universidad de Barcelona.

f) A veces el autor es un organismo o institución. Por ejemplo:

SPSS lnc. (1988). SPSS-X. *User's Guide.* Tercera edición. Chicago: SPSS lnc.

g) En las obras clásicas, cuando interesa especificar el año de la edición original, se puede hacer entre paréntesis después de la referencia consultada. Por ejemplo:

Bernard, C. (1865). *Introducción al estudio de la medicina experimental.* Barcelona: Fontanella, 1976.

h) Cuando existen varias ediciones, la edición concreta se especifica después del título. Ejemplo:

Wittrock, M. (ed.) (1986). *Handbook of Research.* 3ª ed. Nueva York: Macmillan.

i) Cuando son compilaciones *(readings)* se especifica después del nombre del compilador, editor o director. Ejemplo:

Quintana, J. M. (coord.) (1986). *Investigación participativa.* Madrid: Narcea.

j) Cuando se trata de un capítulo de un libro que es una compilación se cita primero el autor del capítulo y luego el compilador. Ejemplo:

Guba, E G. (19&3). Criterios de credibilidad en la investigación *naturalista,* en 1. Gimeno Sacristán y A. Pérez Gómez, *La enseñanza: su teoría y su práctica* (pp. 148-165). Madrid: Akal.

k) Cuando *el* apellido *del* autor es muy corriente se suelen citar los dos apellidos. Por ejemplo:

Fernández Huerta, J. *(1950). Escritura. Didáctica y escala gráfica. Madrid: CSIC*

2. ARTÍCULOS y REVISTAS

Se referencia igual que el libro; en este caso se escribe en cursiva el título de la revista y se especifican: vol., núm. y pp. que ocupa el artículo separadas por un paréntesis. Ejemplo: Mateo, J. (1985). Meta-análisis correlacional sobre estudios de rendimiento escolar en España. *Revista de Investigación Educativa* (RIE), 3, 6, 236-251.

3. OTROS DOCUMENTOS

Cuando se trata de comunicaciones y ponencias presentadas en congresos, seminarios, simposios, conferencias, etc., se especifica autor, título y congreso, añadiendo, si es posible, el mes de celebración. Por ejemplo:

De Miguel, M. (1987). *Paradigmas de investigación educativa.* II Congreso Mundial Vasco, octubre.

4. ORDEN ALFABÉTICO

a) Las referencias bibliográficas deben aparecer ordenadas alfabéticamente por el apellido del autor, o primer autor en caso de que sean varios.

b) Si un autor tiene varias obras, se ordenarán por orden de aparición. Por ejemplo:

De Landsheere, G. (1982). *La investigación experimental en educación.* París: UNESCO.

De Landsheere, G. (19 88). History of Educational Research. En. J. P. Keeves (ed.). *Educational Research, Methodology* & *Measurement. An Intemational Handbook* (pp. 9-16). Oxford: Pergamon.

e) Si son trabajos en colaboración, el orden vendrá indicado por el apellido del segundo autor. Las publicaciones individuales se ponen antes de las obras en colaboración. Ejemplo:

Cronbach, L. J. (1982). *Designing Evaluations of Educational andSocia Programs.* San Francisco: Jossey-Bass;
Cronbach, L. J. Y Suppes, P. (eds.) (1969). *Research for tomorrow's Schools.* Nueva York: Macmillan

d) Si de un mismo autor existen varias referencias, los años se especifican seguidos de letras en orden alfabético *(a, b, c).* Ejemplo:

Freire, P. *(1978a). Pedagogía del oprimido.* Madrid: Siglo XXI;
Freire, P. *(1978b). Pedagogía y acción liberadora.* Madrid: Zero.

3.4 LAS HIPÓTESIS

Concepto y características

Una vez planteado el problema y revisada la literatura, el siguiente paso es preguntarse cuál o cuáles son las soluciones más probables. De entre las soluciones o explicaciones posibles, se eligen las que nos parezcan más plausibles, con el fin de proceder a su contrastación. La explicación o solución posible elegida se denomina *hipótesis.*

Literalmente, la palabra *hipótesis* significa «lo que se pone debajo», lo que se supone. En sentido amplio equivale a suposición o conjetura, expectativa de solución de un problema. Las hipótesis son de gran importancia en el trabajo científico y no es posible avanzar en una investigación si no se comienza por plantear una explicación o solución de la dificultad que la originó. Cuando se formulan en forma de proposiciones, reciben el nombre de *hipótesis* (Arias, 1976).

En la literatura sobre la relación entre hipótesis e investigación encontramos dos posiciones; la de quienes no aceptan que se dé investigación sin hipótesis y la de los que clasifican las investigaciones con o sin hipótesis (Salvador, 1985). Desde este punto de vista las investigaciones sin hipótesis se denominan *exploratorias* y/o *descriptivas,* y las investigaciones con hi12óte¡;i¡;, *explicativas* o *causales.*

En los trabajos no meramente exploratorios o descriptivos, el establecimiento de hipótesis bien formuladas es básico. Es una exigencia metodológica que nos ayudará a orientar toda la investigación y que potenciará su eficacia. En cierta manera, las investigaciones exploratorias, descriptivas o ambas a un tiempo suelen utilizarse para aproximarse a la realidad de los hechos y, a partir de la información obtenida formular con mayor precisión las hipótesis de subsiguientes investigaciones.-.que tratarán de explicarlas.

Las hipótesis científicas tienen también, en general, un sentido de suposición o conjetura. Ary y otros (1987) la definen como «todo enunciado relacional que el investigador formula como respuesta al problema planteado». Para Kerlinger (1985) *es una expresión conjetural de la relación que existe entre dos o más variables.* La

solución tentativa a un problema en forma de proposición comprobable, con la determinación de un grado de probabilidad de certeza o falsedad (Pérez Juste y otros, 1981).

Las hipótesis, para que tengan valor y sean correctas, deben reunir unos criterios de aceptabilidad. Los autores destacan ante todo dos rasgos: que estén bien fundamentadas y que sean contrastables empíricamente. Están bien fundamentadas cuando se hallan insertas en una teoría científica (Bunge, 1976,281); son congruentes con las leyes naturales (Van Dalen y Meyer, 1983, 190) Y están en armonía con otras hipótesis, aunque este aspecto no es esencial (McGuigan, 1976,67). Son contrastables empíricamente cuando se pueden comprobar o rechazar mediante los resultados obtenidos en la realidad. Popper (1971) se niega a aceptar en la ciencia enunciados que no puedan ser sometidos a contraste y pide que la contrastabilidad sea pública. Kerlinger (1985) afirma que si una explicación no puede ser formulada en forma de hipótesis comprobable, se puede considerar como aplicación metafísica y no susceptible de investigación científica.

La hipótesis, poderosas herramientas del investigador que contribuyen al progreso de la ciencia, sirven de puente entre la teoría y la observación; en este sentido, deben ser conceptualmente claras y comprensibles, expresando una relación entre variables; ser una explicación probable del problema o fenómeno a que se refieren, cierto grado de generalización, posibilidad de avanzar predicciones sobre hechos y relaciones aún no observadas. Tiene la función básica de orientar y servir de guía al proceso de investigación; además de desempeñar otras funciones tales como la función de explicación de los fenómenos, de instrumento para el progreso del conocimiento, de nexo entre la teoría y la realidad educativa.

Clasificación de las hipótesis

El tema de la clasificación de las hipótesis se presenta con gran profusión de términos y criterios de clasificación (nivel de generalización, contenido, tipo de variable, etc.) cuya exposición aquí nos llevaría a sobrepasar las pretensiones de este apartado. Vamos a referimos sólo a dos criterios: al origen y al nivel de concreción.

a) *Según el origen, las hipótesis se clasifican en inductivas y deductivas.,*

Las *hipótesis inductivas* surgen de las observaciones o reflexiones sobre la realidad, El investigador formula una hipótesis para hacer una generalización a partir de ciertas relaciones observadas. En sus observaciones advierte tendencias, probables relaciones, y después plantea una hipótesis para explicar dicha relación. Estas hipótesis tienen un carácter más localizado y su poder explicativo es más limitado; ayudan a la resolución de problemas concretos. Se las suele denominar con el nombre de *microhipótesis.* Un ejemplo de hipótesis inductiva sería: «La disposición de los niños para comprometerse en la realización de las tareas de comprensión depende de su percepción del valor de la tarea» (Elliott, 1990).

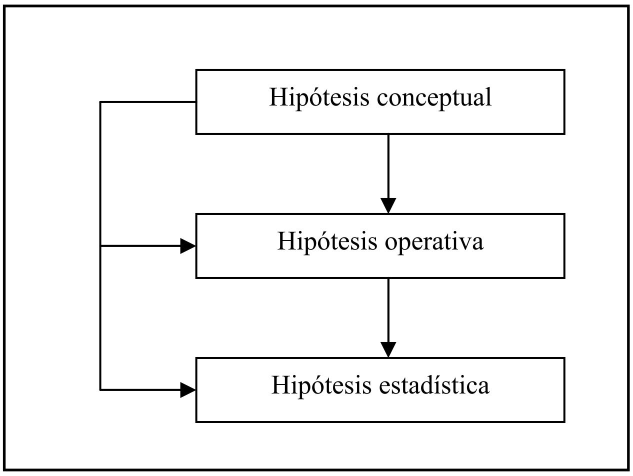
Las *hipótesis deductivas* surgen del campo teórico. Se formulan mediante razonamiento deductivo aplicado a teorías existentes. El proceso sirve para comprobar deducciones implícitas en las teorías. Por lo general, tienen un alcance más amplio que las inductivas y se proponen dar respuesta a problemas teóricos. Un ejemplo de

hipótesis deductiva sería: «La actitud hacia la droga cambiará con la puesta en marcha de programas educativos» (Maciá, 1984).

b) *Según el nivel de concreción, las hipótesis se clasifican en conceptuales, operativas y estadísticas.*

El investigador se ve en la necesidad de salvaguardar la coherencia entre la teoría, la hipótesis, las variables, los datos recogidos y el posterior análisis estadístico que le permitirá confirmar o rechazar la hipótesis. Para ello recurre a definir la hipótesis cada vez con más concreción, pero manteniendo la coherencia interna.

Los niveles de concreción de una hipótesis de investigación podrían representarse mediante el siguiente esquema de flujo:



La *hipótesis conceptual,* denominada también *científica, sustantiva,* expresa una relación conjetural entre dos o más variables definidas de manera abstracta, o bien con relación a una teoría. En la hipótesis «el autoconcepto está relacionado con la autoestima personal», las variables autoconcepto y autoestima tienen una dimensión conceptual, abstracta, referida a una teoría.

En la *hipótesis operativa* las variables se definen operativamente, se indican las operaciones o actividades necesarias para observar, medir o manipular la variable. Una manera de operativizar la hipótesis anterior podría ser formulándola así: «los sujetos que obtienen puntuaciones altas en un cuestionario de autoconcepto, puntuarán también alto en un cuestionario de autoestima personal».

La *hipótesis estadística* expresa la relación de las variables en términos cuantitativos o estadísticos:

$$\bar{X}_a > \bar{X}_b.$$

La hipótesis estadística se enuncia de dos formas: hipótesis nula (Ho) e hipótesis alternativa (H1).

Formulación de hipótesis

Un rasgo fundamental de la hipótesis es que se formule correctamente. Una hipótesis bien planteada ha de ser coherente con el problema y debe especificar los referentes empíricos o indicadores que posibiliten su contrastabilidad.

La formulación formal de la hipótesis hace referencia a la relación establecida entre las variables; según como se formule la relación, las hipótesis se puede enunciar mediante:

a) *Implicación condicional.* Consiste en el enunciado lógico: «Si..., entonces... ». Es decir, si X, entonces Y; si se da la condición X, ocurrirá la consecuencia Y. Si un grupo de alumnos de enseñanzas medias utiliza técnicas de estudio, entonces aumentará su rendimiento escolar. «Si frustración, entonces... agresión».

b) Enunciado preposicional. La forma más habitual, aunque no exclusiva, de formular la hipótesis es en forma de proposición. En forma de enunciado declarativo o afirmación que expresa una relación entre variables. Veamos algunos ejemplos:

- Los alumnos que cursan la asignatura de investigación educativa a través de seminario rendirán más que los que sigan la asignatura a través de clase magistral.

- Las alumnas que optan por seguir carreras de ciencias manifiestan mayor necesidad de éxito que las alumnas que no optan por tales carreras.

- Los alumnos que aprenden a leer por el método global-natural desarrollan más hábitos o estrategias intelectivas que los que utilizan el método silábico o fónico.

Todas estas hipótesis están enunciadas en forma preposicional y relacionan dos o más variables: metodología con rendimiento; carrera con éxito; método con hábitos o estrategias intelectivas.

c) *Hipótesis nula.* Establece que no hay diferencias entre los estadísticos de diferentes muestras extraídas de una misma población. Se denomina así por aceptar que la diferencia entre los estadísticos (medias, varianzas, etc.) de las dos muestras es nula, es decir, las diferencias observadas son atribuibles a oscilaciones del azar. Así, si elegimos dos métodos de lectura para niños de primero de primaria, y planteamos la hipótesis (H) de que el método «A» es superior al método «B», la hipótesis de nulidad se podría formular:

$$H_o: \overline{X}_a = \overline{X}_b$$

Bajo la hipótesis «alternativa» suponemos que las diferencias observadas no pueden ser explicadas por las oscilaciones del azar, es decir, las diferencias entre los estadísticos (medias, varianzas, etc.) de las dos muestras son estadísticamente significativas. En expresión matemática:

$$H_1: \overline{Xa} \neq \overline{X}_b.$$

¿Cómo mejorar la redacción de las hipótesis? Señalamos algunas sugerencias:

1) Formular las hipótesis después de haber revisado las fuentes.
2) Enunciar varias hipótesis como posibles soluciones del problema que se estudia.
3) Redactadas en forma de proposición afirmativa. En forma interrogativa sólo si después de la revisión de las fuentes no puede deducirse dirección alguna.
4) El enunciado ha de ser correcto y operativo, sin perder claridad.
5) Que la hipótesis se pueda contrastar, confirmar o rechazar.
6) Definir con claridad cada término de la hipótesis.
7) Recurrir a hipótesis múltiples que ayuden a redefinir el problema.
8) Evitar juicios de valor: «La técnica de la enseñanza de las matemáticas con diapositivas es "mejor" que la exposición magistral».

Contrastración de hipótesis

El investigador suele formular por lo general varias conjeturas como posibles soluciones al problema. Selecciona entre ellas la que considera mejor y después la somete a contrastación para observar si se confirma o si debe rechazarla.

Una forma de validar las hipótesis es intentar contrastar hipótesis bien fundamentadas o bien tratar de fundamentar teóricamente las hipótesis confirmadas, si bien es posible que puedan plantearse hipótesis bien fundamentadas pero que no sean contrastables o que puedan encontrarse hechos empíricos cuyas hipótesis explicativas no encuentren fundamentación teórica.

La tarea del investigador consiste en contrastar la hipótesis en el sentido de confirmarla, y no en probarla y verificarla, como sostienen algunos autores. Verificar implica probar que es verdad y esto no es lo que se pretende al confirmar la hipótesis. Según Van Dalen y Meyer (1983, 18) «el hecho de obtener elementos empíricos que concuerden con las consecuencias que se derivan de una hipótesis, no implica probar una hipótesis, sino sólo determinar que ella posee cierto grado de probabilidad».

Conviene recordar que ninguna hipótesis se contrasta aislada del contexto teórico en el que están implicados aspectos como selección de unas variables, una muestra, un momento y un lugar determinado.

Retornando el esquema de flujo señalado anteriormente, la contrastación de las hipótesis seguiría estos pasos:

a) El investigador parte de la hipótesis conceptual (sustantiva) donde se expresa una relación conjetural.
b) La hipótesis conceptual se transforma en operativa. Se define en función de cómo se observan, miden o manipulan las variables.
c) La hipótesis opérate va se transforma en hipótesis estadística en términos cuantitativos y estadísticos.
d) La hipótesis estadística se prueba con la hipótesis nula (proposición estadística que afirma que no hay relación entre las variables):

$$H_0: \overline{X}_a = \overline{X}_b \text{ o } H_1: \overline{X}_a - \overline{X}_b = 0$$

La hipótesis nula establece que no existe diferencia entre la proposición hipotética y la que se encuentra presente en la población; es una proposición de no diferencia entre los resultados empíricos y los teóricos. Estrictamente hablando no podemos probar la hipótesis nula (H_o), ni tampoco probar directamente la hipótesis alternativa (H_1).

e) A la hipótesis nula la podemos aceptar o rechazar. Si la H_1 podemos rechazar, podemos aceptar la hipótesis alternativa, que afirma que

$$\overline{Xa} \neq \overline{X_b},$$

¿En qué sentido? El sentido de la H_1 es «bidireccional»; $\bar{X}a$ puede ser mayor o

menor que $\bar{X}b$. Si los datos apuntan en una dirección, la hipótesis científica se convierte en «direccional» o «unilateral».

3.5 ESTUDIO DE LAS VARIABLES

El estudio de las variables ocupa un lugar relevante en el proceso de la investigación educativa, Al formular las hipótesis, de ordinario el investigador identifica y reconoce las variables que entran en el estudio. La hipótesis como proposición conjetural relaciona conceptos formulados en términos de variables, estableciendo entre ellos algún tipo de nexo. Las variables suelen derivar de otros conceptos de mayor nivel de abstracción llamados *constructos,* que cuando los definimos en términos más concretos y observables los denominamos *variables.*

Concepto de variable

Una variable es una característica o atributo que puede tomar diferentes valores o expresarse en categorías. Así, decimos que son variables la temperatura (18°C, 23°C), el coeficiente intelectual (95, 102), el método de enseñanza (tradicional, activo), el sexo (varón, hembra). Se entiende por *variable* las cualidades o aspectos en los que difieren los fenómenos o individuos entre sí. El concepto *variable* deriva del concepto matemático de variable como oposición a *constante,* característica que sólo puede adoptar un único valor o categoría para todos los individuos.

Algunos autores definen la variable como «todo fenómeno considerado en función de una de sus características que al manifestarse puede tomar distintos valores, de acuerdo con un sistema definido de clasificación» (De la Orden, 1975, 225). Kerlinger (1985) afirma que variable «es un símbolo al que se le asignan numerales o valores». En un sentido general, Hayman (1979, 68) define variable como «algo que puede cambiar, cuantitativa o cualitativamente», y Freeman (1970) como «características observables de algo que son susceptibles de adoptar distintos valores o de ser expresadas en varias categorías». Finalmente, para Carlsmith y otros (1976,10) las variables serán entendidas como «cualquier atributo que puede asumir valores diferentes entre los miembros de una clase de sujetos o sucesos, pero que tiene un solo valor para cualquier miembro de esa clase en un momento dado».

Constructos y variables

En investigación educativa empírica los fenómenos que se estudian se expresan en conceptas científicos que reciben el nombre de *constructos hipotéticos. San* entidades abstractas supuestas, bien definidas y articuladas, que consideramos que existen, aunque *no.* sean estrictamente observables y que sirven para explicar determinadas fenómenos (Hayman, 1979). *Par* ejemplo., la inteligencia, la motivación, la capacidad de aprendizaje, san constructos.

Definición de variables Los constructos y variables se hallan relacionados de tal forma, que constituyen, en cierta manera, las das caras de la misma moneda. Los constructos hipotéticos se consideran como variables latentes, no observables directamente. Para estudiar los constructos lo que se hace es analizar unas variables observables que funcionan como «indicadores» del fenómeno. no observable (Bisquerra, 1987). Así, *par* ejemplo., del constructo «inteligencia» (aptitud para enfrentarse a situaciones nuevas con rapidez y éxito.) se puede pasar a la variable «inteligencia» *cama* puntuación obtenida en un test; del *constructo* «capacidad de aprendizaje» a la variable «rendimiento escolar» expresada en las calificaciones escalares.

Una de los problemas que encontramos al diseñar una investigación es *cómo* definir correctamente los constructos en términos de variables. Una manera de superar esta dificultad es definir las variables en términos operativas. Consiste en el pasa que hemos señalada en el apartado anterior al hablar de la hipótesis conceptual y de la operativa. Las variables se pueden definir de varias formas:

De forma conceptual o *constitutiva;* es decir, definiendo en qué consiste la naturaleza de la variable mediante conceptos; por ejemplo., la inteligencia definida *como* «aptitud para enfrentarse a situaciones nuevas *can* rapidez y éxito», es una definición constitutiva.

De forma operativa. Consiste en definir las variables describiendo las operaciones *o* actividades que han de realizarse para medirla o manipularla. Se convierte en una serie de instrucciones que el investigador debe tener presentes para medirla o manipularla. El investigador ha de traducir las variables a definiciones observables, cuantificables y medibles. Así, *por* ejemplo., la variable «nivel de instrucción» puede definirse *como* «grado de instrucción que posee la persona» y la variable «fluidez verbal», *como* «cantidad de palabras escritas en un tiempo dado» y la variable «profesor directivo», *como* «profesor que da instrucciones y establece relaciones formales *con* los alumnos».

Las operaciones *o* actividad que realiza el investigador para pasar del campo teórico, los constructos, al campo de la realidad observable, variable operativa, se designa *como operativización de la variable.* Consiste en sustituir unas variables por otras más concretas que sean representativas de aquéllas. Kerlinger (1985) califica este tipo de definición de invento maravilloso que establece un puente entre los conceptos y las observaciones y actitudes reales. Asigna significado a la variable especificando las actividades u operaciones necesarias para medirla. Insistimos, pues, en que una definición operativa se basa en características observables de lo que se está definiendo.

Una misma variable puede ser operativizada de diversas formas. Por ejemplo., la

inteligencia puede medirse con varios tests: Raven, Terman, Wechsler, etc. El instrumento que se utiliza condiciona los datos; de ahí la importancia de garantizar la validez de la operativización.

Mediante la operativización de variables se pretende pasar de variables conceptuales a otras operativas, y de éstas a indicadores observables. Esta transformación tiene su importancia en la realización de una investigación empírica.

¿Cuál es el procedimiento? Sierra Bravo (1987, 110) recoge la aportación de otros autores y distingue las siguientes fases:

l. Se enuncia o define la variable.
2. Se deducen sus dimensiones o aspectos principales.
3. Se buscan indicadores o circunstancia empíricas concretas de cada dimensión.

Un ejemplo que resume estas tres fases sería: seleccionar la variable «clase social» y tomar como variable intermedia el «nivel económico». Indicadores de éste serían las fuentes de renta: sueldos, fincas, intereses de capitales, etc.

4. Se construyen los índices.

No todos los indicadores poseen la misma importancia, de ahí la necesidad de construir un índice que agrupe en una medida común todos los indicadores referentes a una dimensión mediante la asignación de un peso o valor a cada uno de ellos de acuerdo con su importancia.

Las escalas de medida

Las escalas de medida se refieren al sistema de medida que el investigador aplica a las variables. La naturaleza de las variables determina el tipo de escala.

Se señalan cuatro niveles de medida; cada uno posee sus propias características y, en cada uno de ellos, están o no permitidas determinadas operaciones aritméticas. El tipo de escala depende de las operaciones empíricas que permite efectuar. En orden de precisión creciente, las cuatro escalas o niveles de medida son: la nominal, la ordinal, la de intervalos y la de razón.

Escala nominal. Se aplica a las variables cualitativas. Permite clasificar los objetos e individuos según sean iguales o no con respecto a una característica. Determina empíricamente la operación de igualdad o desigualdad (=). Ejemplos, color del cabello (castaño, moreno, rubio...), sexo, las líneas de autobús (4, 14, 24...), etc. Estadísticos aplicables a la variable medida con esta escala son la frecuencia, la moda, el coeficiente de contingencia, etc.

Escala ordinal. Clasifica los objetos e individuos según el orden que ocupan con respecto a una característica. Permite realizar las operaciones mayor que (>) y menor que (<), además de la de equivalencia (=). Ejemplos, clase social, orden de llegada a la meta, los centiles de un test, dureza de los minerales. Estadísticos aplicables a la variable medida con esta escala son la mediana, los centiles, la correlación ordinal, etc.

Escala de intervalos. Permite medir los objetos e individuos indicando si son superiores o no a otros con respecto a una característica e indicar la distancia entre ellos. Es decir, comprobar la igualdad de diferencias entre los valores. Los valores numéricos son cuantitativos, se pueden agrupar en intervalos de unidad constante y se pueden realizar las operaciones aritméticas fundamentales. Las escalas de intervalo poseen un cero arbitrario, pero no absoluto. Ejemplo, la temperatura, las puntuaciones de un test, el año del calendario. Estadísticos aplicables a la variable medida con esta escala son la desviación típica, la correlación, etc.

Escala de razón o proporcional. Son escalas que añaden a las de intervalo la existencia del cero absoluto. El cero significa ausencia total de la característica que hay 'que medir. Se considera la escala más perfecta. Permite determinar la igualdad de razones, calcular la razón significativa entre dos valores de la escala. Ejemplos, el peso, la altura, la edad, etc. Esta escala es difícil de utilizar en educación, debido a su condición de cero absoluto. Estadísticos aplicables a la variable medida con esta escala son la media geométrica, el coeficiente de variación, etc.

Clasificación de las variables

Los autores suelen mantener criterios y utilizar terminología distinta a la hora de clasificar las variables (cfr. Tuckman, 1971; McGuigan, 1976; Pérez Juste, 1981; Kerlinger, 1985; etc.). Estas suelen clasificarse según distintos criterios:

a) SEGÚN EL PUNTO DE VISTA TEÓRICO-EXPLICATIVO

Variables estímulo. Cualquier condición externa natural o social que pueda afectar a la conducta del individuo. Ejemplo, la temperatura ambiental.

Variables respuesta. Las que se manifiestan en la conducta del individuo como efecto de una variable estímulo. Ejemplo, llorar viendo una película.

Variables organísmicas. Cualquier característica del individuo que pueda mediar o intervenir entre las variables anteriores. Ejemplo, el sexo, la motivación. '

b) SEGÚN LA NATURALEZA DE LAS VARIABLES

Cualitativa. También llamada *atributo* o *categórica.* Es una característica que se expresa en categorías, debido a que por su naturaleza no es cuantificable. Según el número de categorías o modalidades que toma se divide en *dicotómica* y *politómica.*

La *dicotómica* admite o se expresa en dos categorías, ejemplo, el sexo: varón o hembra; tipo de centro: estatal o privado; ítems de verdadero-falso. *La politómica* admite o se expresa en más de dos categorías; ejemplo, la clase social: superior, media o baja; el nivel de estudios.

El número de categorías en que se expresa una variable es, en general, una decisión arbitraria; no obstante, su clasificación ha de cumplir algunas propiedades:

- Homogeneidad: las categorías deben guardar relación lógica con la variable y entre sí; deben estar bien definidas.
- Inclusión: las categorías deben ser exhaustivas, incluir a todos los sujetos.
- Utilidad: cada categoría debe ser significativa respecto al problema de investigación.
- Exclusión: las categorías deben ser mutuamente excluyentes; cada observación sólo podrá ser clasificada en una categoría.

Cuantitativa. Es la característica que se puede medir y expresar en valores numéricos referidos a una unidad de medida. Ejemplo, la estatura, el rendimiento escolar. Esta, a su vez, se subdivide en:

- *Cuantitativa discreta.* Es la variable que por su naturaleza intrínseca sólo puede tomar determinados valores enteros. Ejemplo, número de hijos (0, 1, 2, 3,..) número de alumnos en clase.

- *Cuantitativa continua.* Es la variable que puede tomar cualquier valor entero o fraccionario dentro de un continuum. Ejemplo, el peso, el rendimiento académico (6; 6,3; 6,05...).

c) SEGÚN EL CRITERIO METODOLÓGICO

Es el criterio más importante de clasificación en investigación empírica. Las variables en función del papel que se les asigna en la investigación se clasifican en independiente, dependiente e interviniente.

Variable independiente (VI). Es la característica que el investigador observa o manipula deliberadamente para conocer su relación con la variable dependiente .. La VI es la situación antecedente de un efecto; responde a la idea de causa, si bien en educación resulta más propio hablar de relación. A veces se le denomina con los nombres de *estímulo, experimental* o *tratamiento.* En la formulación de la hipótesis se reconoce a la VI como el antecedente y a la VD como el consecuente. En los experimentos en sentido estricto, la VI es manipulada por el investigador, convirtiéndose, en terminología de Kerlinger, en una *variable activa.* Para Ary y otros (1987) las variables independientes, según sean manipulables o no, se clasifican en activas y asignadas o atributivas.

- *Variable activa.* Es la variable independiente susceptible de manipulación directa y deliberadamente. Ejemplo, el método de enseñanza, un programa de instrucción, etc.

- *Variable atributiva o asignada.* Es la variable independiente que no es posible manipular activamente. Ejemplo, la actitud, la edad, el sexo, etc.

Variable dependiente (VD). Es la característica que aparece o cambia cuando el investigador aplica, suprime o modifica la variable independiente. Suele denominarse

criterio y corresponde a la idea de efecto producido por los cambios de la VI. En educación la VD por excelencia suele ser el rendimiento escolar.

Que una variable sea independiente o dependiente dependerá del uso o función que se le asigne en la hipótesis. Así, por ejemplo, si seleccionamos las variables «capacidad para resolver un *puzzle*» e «inteligencia», bien podríamos diseñar una investigación que permita determinar:

a) la influencia de la inteligencia (VI) en la capacidad para resolver *puzzles* (VD);
b) la influencia de la capacidad para resolver *puzzles* (VI) en el desarrollo de la inteligencia (VD).

Variable interviniente. Son las características ajenas al experimento que influyen en los resultados o pueden desvirtuarlos. Hacen referencia a las disposiciones conductuales y ambientales que afectan a los resultados. Kerlinger (1985) las denomina *intermedias* por su condición de mediación. Dentro de esta modalidad entrarían la motivación, rasgos de personalidad, condiciones ambientales, etc.

En los fenómenos educativos se hallan implicadas múltiples variables. Una respuesta, un efecto..., no depende sólo de una única variable, es resultado de la interrelación de múltiples variables independientes, generalmente en interacción.

Cuando queremos conocer la influencia de una variable sobre un efecto o respuesta, se debe procurar controlar el resto de variables que no deseamos que intervengan en la investigación; en este caso nos encontramos con un tipo de variables: las variables extrañas (Pérez Juste y otros, 1981, 180). Son aquellas variables ajenas al proceso de investigación pero que actúan asociadas a las VI, de modo que los resultados apreciados en la VD quedan contaminados. Se definen, por exclusión, como las que no son ni la variable VI ni la VD.

El investigador debe controlar el efecto que las variables extrañas generan en la variable dependiente, o bien estimarlo, y actuar del mismo modo y en el mismo grado sobre todos los grupos tomados en la investigación.

En una investigación de carácter empírico son muchas las variables extrañas que se controlan: edad, sexo, nivel socio cultural, etc. Fox (1981) incluye bajo el epígrafe de variables extrañas a todas las variables que no sean independiente o dependiente. Tous (1978), sin embargo, habla de variables intermediarias, diciendo que son procesos que se hallan entre la variable independiente y la dependiente y que no pueden observarse directamente. Bartolomé (983), por su parte, reserva para el concepto de variable extraña aquellas que son directamente observables, proceden del medio, de los sujetos, bien del propio investigador, y pueden ejercer influencia la investigación modificando sus resultados.

3.6 TÉCNICAS DE MUESTREO

El propósito de la investigación empírica es desarrollar generalizaciones útiles para la educación y los procesos de enseñanza y aprendizaje. Sin embargo, debido a las limitaciones de tipo práctico que presentan los recursos de investigación, el investigador, usualmente, recurre al estudio de muestras extraídas de *poblaciones* a las que quiere generalizar los resultados.

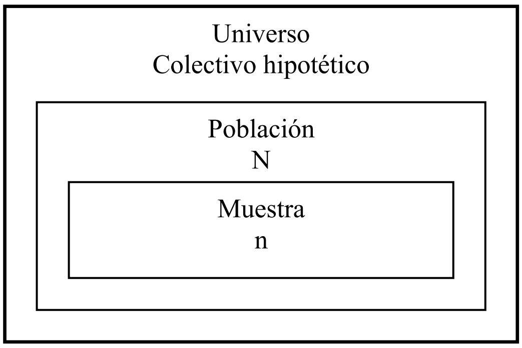
Población y muestra

Una acción importante del investigador es determinar los sujetos con quienes se va a llevar a cabo el estudio; lo que hace necesario delimitar el ámbito de la investigación definiendo una población.

En la práctica resulta difícil que una investigación se lleve a cabo con todos los casos que componen la población por razones de tiempo, dinero, accesibilidad a los sujetos. Esta dificultad se soslaya mediante la selección de muestras representativas de la población.

Un conjunto finito de unidades puede extraerse de un conjunto más amplio y también finito de unidades. Este conjunto parcial, extraído de otro, se constituye en una muestra del conjunto mayor, al que suele llamarse *colectivo* o también *universo.*

Los distintos conjuntos se pueden representar gráficamente según este esquema:



Conceptos fundamentales

/ *Universo.* Designa a todos los posibles sujetos o medidas de cierto tipo.

/ *Población.* Es el conjunto de todos los individuos (objetos, personas, eventos, etc.) en los que se desea estudiar el fenómeno. Estos deben reunir las características de lo que es objeto de estudio.

/ *Muestra.* Conjunto de casos extraídos de una población, seleccionados por algún método de muestreo. Se considera muestra grande cuando *n* es = 30 o mayor.

/ *Muestra*: Cada uno de los elementos que componen la muestra y de los que se obtiene la información. Los individuos pueden ser personas, objetos o acontecimientos.

Muestreo

Lo esencial del muestreo es identificar la población que estará representada en el estudio. La extracción de una muestra a partir de la población se conoce con el nombre de *muestreo.*

Fox (1981, 367-369) señala cinco etapas en el proceso de muestreo:

1) definición o selección del universo o especificación de los posibles sujetos o elementos de un determinado tipo;
2) determinación de la población o parte de ella a la que el investigador tiene acceso;
3) selección de la muestra invitada o conjunto de elementos de la población a los que se pide que participen en la investigación;
4) muestra aceptante o parte de la muestra invitada que acepta participar;
5) muestra productora de datos: la parte que aceptó y que realmente produce datos.

El mismo autor propone un ciclo de muestreo que se esquematiza en la figura 3.5.

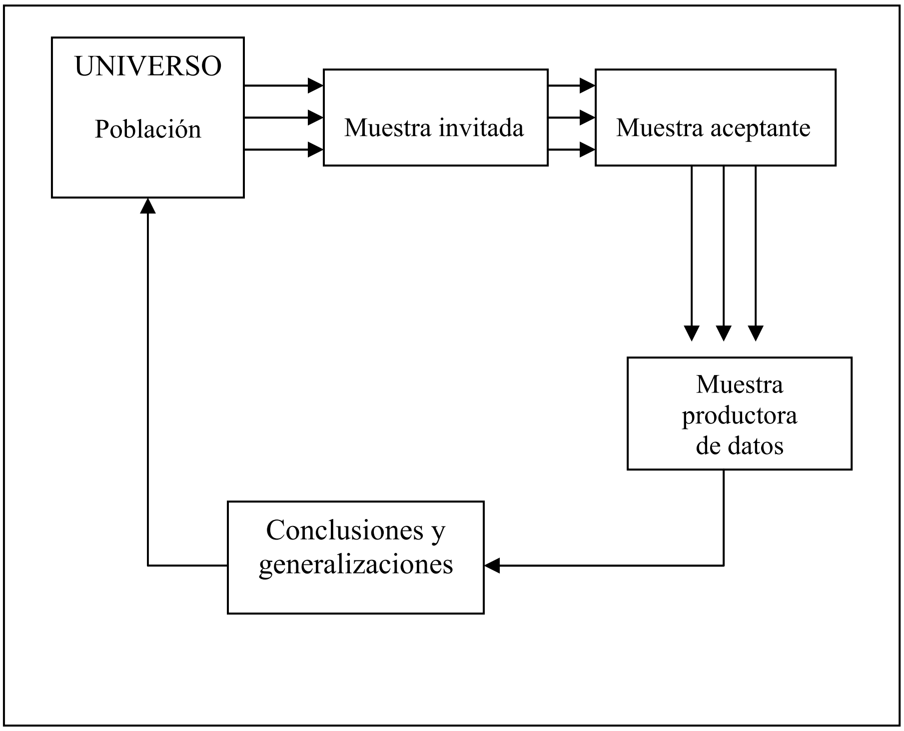


FIG.3.5 Ciclo de muestreo (Fox, 1981, 369)

Entre las ventajas que proporciona el muestreo suelen señalarse: *a)* ahorro de tiempo en la realización de la investigación; *b)* reducción de costos, y e) posibilidad de mayor profundidad y exactitud en los resultados.

Entre los inconvenientes se señalan: *a)* la dificultad de utilización de la técnica de

muestreo; una muestra mal seleccionada distorsiona los resultados; *b)* las limitaciones propias del tipo de muestreo, y c) tener que extraer una muestra de poblaciones que poseen pocos individuos con la característica que hay que estudiar.

Tipos de muestreo

Si bien no existe una única manera de clasificar los tipos de muestreo, es corriente divididos en probabilísticos y no probabilísticos.

a) PROBABILÍSTICOS

En los muestreos probabilísticos se cumple el principio de equiprobabilidad; que quiere decir que todos los individuos tienen la misma probabilidad de salir elegidos en una muestra. Dentro de los probabilísticos existen varias modalidades:

Muestreo aleatorio simple. Es la modalidad de muestreo más conocida y que alcanza mayor rigor científico. Garantiza la misma probabilidad de ser elegido a cada elemento de la población y la independencia de selección de cualquier otro. Puede realizarse con o sin reposición.

Los pasos que se realizan para seleccionar una muestra aleatoria suelen ser: *a)* definir la población y confeccionar una lista de todos los individuos y asignarles números consecutivos de 1 hasta *n; b)* la unidad de base de la muestra debe ser la misma; c) concretar el tamaño de la muestra, y *d)* extraer al azar los elementos.

La muestra quedará formada por los *n* individuos obtenidos mediante sorteo del total de la población. Entre los procedimientos que permiten extraer los individuos en el muestreo aleatorio simple están: *a)* las tablas de números aleatorios que incluyen los textos de estadística; *b)* los clásicos sistemas de lotería; y c) otros procedimientos de extracción al azar.

Muestreo sistemático. Es una variante del anterior. Primero se calcula 1 mediante la fórmula:

$$I = \frac{N \text{ (población)}}{n \text{ (muestra)}}$$

Luego se elige al azar un número menor o igual que 1, es decir, que cumpla la condición

$$1 \leq a \leq I.$$

Por último se seleccionan los números, que serán:

a, a + I, a + 2I, a + 3I, etc.

Si en una facultad hay 3000 alumnas y se quiere realizar un estudia can 200, el intervalo. será 3000/200 = 15. Utilizando. la tabla de números aleatorias, cada 15 números se seleccionan las sujetas que formaran la muestra partiendo. de a.

Muestreo estratificado. Este muestreo se utiliza cuando la población está constituida en estratos, conjuntos de la población con homogeneidad con respecta a la característica que se estudia. Dentro de cada estrato se puede aplicar el muestrea aleatorio o el sistemático..

Para la obtención de una muestra estratificada se siguen los siguientes pasos: *a)* se divide la pab1ación en estratos; *b)* de cada estrato se extrae una muestra por algún procedimiento de muestreo; c) el número de individuos de cada estrato se puede decidir por afijación simple (la muestra total se divide en partes iguales); afijación proporcional (se tiene en cuenta la proporción de individuas en cada estrato.); afijación óptima (además de la proporción de cada estrato. se tiene presente la varianza de los datos, y en las estratos donde ésta sea grande, se tomará una muestra de mayor tamaño.); y d) la suma de las muestras de cada estrato forman la muestra total *n.* Dentro del muestreo estratificado suelen distinguirse:

- *Muestreo estratificado proporcional.* Cada estrato queda representado en la muestra en proporción exacta a su frecuencia en la población total. Si el 15 % de la población son estudiantes universitarios, el 15 % de la muestra se extraerá de este estrato..

- *Muestreo estratificado constante.* La muestra se obtiene seleccionando un número igual de individuos de cada estrato.. El 50 % de hombres y el 50 % de mujeres.

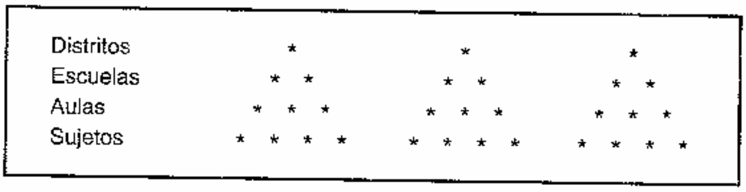
Muestreo por conglomerados o grupos. Este tipo de muestreo se utiliza cuando *los* individuos de la población constituyen grupos naturales o conglomerados (distrito., centro escolar, etc.). La unidad muestral es el conglomerado. *(cluster) a* grupo y no los individuos *como* en los anteriores. La selección aleatoria se aplica a los conglomerados y no a los individuos que los componen. El muestreo por conglomerados o grupas sigue estos pasos:

a) La población se divide en grupas o. cang1omerados;
b) se seleccionan aleatoriamente los conglomerados;
c) los sujetos de *los* conglomerados constituyen la muestra.

Este muestreo ofrece la ventaja de que no es necesario identificar a todos las individuos de la población para seleccionarlos a1eatariamente, sino sólo los individuos de los conglomerados se1eccianados, sin que importe el resto..

Muestreo polietápico. Sigue una secuencia de etapas de selección de unidades muestrales de mayor rango a otros de menar, hasta llegar a los individuos que constituyen la muestra (distrito escolar, escuelas, aulas, alumnos). Este tipo de muestreo sólo necesita conocer los individuos que integran los conglomerados de la última etapa. Además, en cada etapa puede aplicarse un muestreo aleatorio.

b) NO PROBABILÍSTICOS



Estas técnicas no utilizan el criterio de equiprobabilidad, sino que siguen otros criterios, procurando que la muestra obtenida sea lo más representativa posible. Los más conocidos son:

Muestreo accidental o casual. El criterio de selección de los individuos depende de la posibilidad de acceder a ellos. Es frecuente utilizar sujetos que las condiciones nos permiten. Por ejemplo, entrevistar a la salida de un metro, o a las personas que pasan por una calle.

Muestreo intencional u opinático. Se eligen los individuos que se estima que son representativos o típicos de la población. Se sigue un criterio establecido por el experto o investigador. Se seleccionan los sujetos que se estima que pueden facilitar la información necesaria.

Muestreo por cuotas. Consiste en fijar unas «cuotas». Cada cuota consiste en un número de individuos que reúnen unas determinadas condiciones. La selección de las «cuotas» suele hacerse mediante «rutas» o «itinerarios».

Representatividad de la muestra

El muestreo tiene sentido en cuanto garantiza que las características que se quieren observar en la población queden reflejadas en la muestra. Preservar la representatividad es el atributo más importante que debe reunir el muestreo, lo que nos permitirá generalizar a la población los resultados obtenidos en la muestra.

Fox (1981, 373) señala que para conseguir la representatividad se requieren tres condiciones:

a) saber qué características (variables) están relacionadas con el problema;
b) capacidad para medir esas características (variables);
c) poseer datos de la población sobre las características para utilizarlos como base de comparación.

Según el mismo autor, si no se cumple alguna de estas condiciones para alguna de las características, se pierde la capacidad de buscar deliberadamente la representatividad en cuanto a ella.

La utilización de los procedimientos aleatorios en el muestreo ayudan a garantizar

la representatividad de la muestra, pero ello no es garantía total para que estemos seguros de que la muestra al azar es representativa de la población de que se ha extraído (Kerlinger, 1985).

Cuando una muestra seleccionada al azar resulta sesgada, el investigador tiene dos opciones: *a)* proseguir el estudio utilizando la muestra no representativa; o *b)* seguir seleccionando otros elementos de la población aunque el estudio tenga una muestra mayor de la que se buscaba inicialmente.

¿Cómo se puede contrastar la representatividad de la muestra? Recurriendo a los tres tipos de muestra:

a)invitada (conjunto de individuos «invitados» de la población);
b) aceptante (conjunto de individuos que aceptan participar);
e) productora de datos (muestra real del estudio).

Si se tienen datos de la población, se pueden comparar con ella la muestra invitada y la aceptante. Se puede utilizar algún método adecuado para contrastar la significación y determinar si las muestras son representativas o no de la población, como la chi-cuadrado u otras.

Si el contraste indica que no hay diferencias se puede admitir la representatividad de la muestra; en caso contrario se puede recurrir a las dos opciones señaladas anteriormente: trabajar con la muestra no representativa o seleccionar más elementos, hecho que puede ayudamos a conseguir la representatividad.

Si no se tienen datos de la población, no se puede realizar el contraste de la muestra invitada y aceptante con respecto a la población. Entonces se procede a contrastar la muestra aceptante con la invitada.

¿ Qué pérdida suele tolerarse desde la selección de la muestra hasta llegar a la real o productora de datos? No existe una única respuesta. La pérdida del 25 % debe preocupar; cuando sea inferior al 50 % «se debe leer y escribir con cuidado»; y cuando la proporción es menor del 40 % no se deberían dar a conocer los datos, ni considerarlos como conclusiones válidas. Son útiles como estudios pilotos, pero no se pueden aceptar sin hacer un estudio posterior más exhaustivo.

Tamaño de la muestra

Cuando realizamos investigación con poblaciones, una de las primeras preguntas que es preciso plantear es qué número de sujetos o tamaño de la muestra se necesita. Este tiene que ser compatible con la representación de la población. Compaginar el tamaño de la muestra y la representatividad no es tarea fácil. Fox (l981, 396) señala que es antes la representatividad que el tamaño de la muestra. Ante todo la muestra debe ser lo suficientemente grande que garantice la representatividad.

Para seleccionar el tamaño de la muestra se utilizan varios procedimientos: curvas, fórmulas y tablas. De ordinario, su estimación requiere que se definan previamente el nivel de confianza y el error de estimación.

Bugeda (1975) ofrece tablas que permiten determinar el tamaño de la muestra y el nivel de confianza correspondiente; hecho que facilita la economía de esfuerzos en la utilización innecesaria de muestras grandes.

Para estimar el tamaño de la muestra necesitamos conocer el nivel de confianza que queremos que alcancen los datos; frecuentemente se opta por un nivel de confianza del 95 o 99,7%, y un error de estimación máximo de un 5%. Según se trate de poblaciones infinitas o finitas, la estimación del tamaño de la muestra varía:

a) Determinación del tamaño de la muestra para poblaciones infinitas (más de 100 000 individuos). Si deseamos estimar una proporción de una población a partir de una muestra, pueden presentarse dos situaciones:

a.1) Que conozcamos la proporción de sujetos que posee la característica a través de estudios previos. En este caso se aplica la fórmula:

$$n = \frac{Z^2_{\alpha} \cdot p \cdot q}{e^2}$$

n = número de elementos que debe poseer la muestra
ex = riesgo o nivel de significación
Za = puntuación correspondiente al riesgo ex que se haya elegido. Por ejemplo, para
un riesgo ex = 0,05 *(Za* =1,96)
P = % estimado
q = 100-p
e = error permitido.

Veamos el siguiente ejemplo: ¿Cuántos sujetos debe tener una muestra si se desea estimar la proporción de sujetos que opinan de una determinada manera al nivel de confianza del 99,7% (3 **σ**) Y un error de estimación de 5%, sabiendo que p es el 25%?

$$n = \frac{3^2 \cdot 25 \cdot 75}{5^2} = 675 \text{ sujetos.}$$

a.2) O que desconozcamos la proporción de sujetos. En tal caso se toma p = 50% Y q = 50%
¿Qué número de sujetos debe tener una muestra si se desea estimar qué proporción de sujetos poseen una característica al nivel de confianza del 99,7% (3**σ**) y un error de estimación admitido de 2%?

$$n = \frac{3^2 \cdot 50 \cdot 50}{2^2} = 5625 \text{ sujetos.}$$

En el caso de que la investigación requiera estimar una media podemos acudir a esta fórmula:

$$n = \frac{Z^2_{\alpha} \cdot \sigma^2}{e^2}.$$

b) Determinación del tamaño de la muestra para poblaciones finitas (menos de 100000 individuos). Se desea calcular el número de sujetos necesarios para realizar investigación en una provincia en la que terminan octavo curso 60 000 alumnos. Se establece el nivel de confianza en 20', es decir, 95,5%, y el margen de error, en e13%. Otros estudios anteriores indican que el 50% de los alumnos de octavo curso de enseñanza elemental optaban por el bachillerato.

$$n = \frac{Z\alpha^2 \cdot p \cdot q \cdot N}{e^2 (N\text{-}l) + Z\,\alpha^2\, p.\,q}$$

Sustituyendo las letras por sus valores:

$$n = \frac{4 \cdot 50 \cdot 50 \cdot 60\,000}{9\,(60000 - 1) + 4 \cdot 50 \cdot 50} = 1091 \text{ sujetos.}$$

En caso de que la investigación requiera estimar una media podemos acudir a esta fórmula:

$$n = \frac{N \cdot Z_\alpha^2 \cdot \sigma^2}{Z_\alpha^2 \cdot \sigma^2 + e^2 \cdot (N-1)}$$

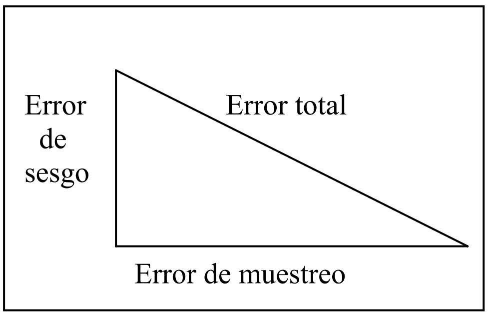
Cuando sea necesario contrastar medias o proporciones puede acudirse a libros de estadística especializados en el tema.

Error muestral

Al tomar una muestra de una población es frecuente que los resultados obtenidos de la muestra no sean exactamente los valores reales de la población. El error muestral se define como la diferencia entre el parámetro de una población y el estadístico de una muestra. Los errores de muestreo surgen al estudiar una muestra en lugar de la población total.

Los errores muestrales son debidos a dos causas principales: *a)* error de sesgo (la muestra no es representativa de la población) y *b)* error aleatorio (debido al azar).

Estos dos tipos de errores se relacionan entre sí mediante la expresión (error total)^2 = (error por sesgo)^2 + (error de muestreo)^2



Para estimar el error muestral existen varias técnicas denominadas de *submuestreo (subsample)* o *remuestreo (resample)* que se basan en dividir la muestra total en dos o más submuestras aleatorias y a partir de ellas estimar los estadísticos y determinar las propiedades del estimador (Wilson, 1988).

De entre las técnicas de «submuestreo» señalaremos algunas: *a)* la técnica de replicación independiente (la muestra se divide en submuestras independientes); *b)* la técnica del *jakknife* (navaja) (se corta la muestra en varias partes no mutuamente exc1uyentes, es decir, con reposición), y c) replicación repetida balanceada (Ross, 1988).

3.7 DIVERSIDAD METODOLÓGICA

Una dimensión importante del proceso de investigación es la metodología que hay que utilizar, que consiste en la manera de llevar a cabo la investigación o modo de enfocar los problemas y buscar las respuestas; en este sentido la metodología se interesa más por el proceso que por los resultados.

El conocimiento de la metodología es de gran utilidad para el investigador, ya que le permite seleccionar la perspectiva metodológica adecuada al problema planteado. Ésta dependerá de la finalidad y objetivos que el investigador se proponga. En este apartado vamos a describir y precisar los atributos que constituyen y singularizan la naturaleza de los distintos enfoques o perspectivas metodológicas, sin entrar en su análisis y valoración crítica. Antes precisaremos algunos conceptos propios de este tema.

Metodología, método y técnica: aspectos diferenciales

En el ámbito de la investigación educativa los términos *metodología, método* y *técnica* son conceptos que se caracterizan por su ambigüedad, aparecen con diferentes significados y se les atribuyen diversos usos y fines; es frecuente usarlos como sinónimos cuando en realidad no lo son. Conceptualizarlos y describir su naturaleza no es fácil y puede resultar arbitrario. No obstante, consideramos necesario precisar la terminología y, al menos a nivel conceptual, clarificar su significado.

Metodología. El término *metodología* se caracteriza por su naturaleza multidimensional, lo que hace difícil su concepción y precisión. En los textos se utiliza con diversos sentidos, que generalmente se refieren a los componentes -teórico, morfológico y técnico- que constituyen su naturaleza. Varias son las acepciones que aparecen en los textos:

a) Como «la lógica de la investigación». En sentido general, en las ciencias sociales. por *metodología* se entiende la manera de realizar la investigación; el modo de enfocar los problemas y de buscar las respuestas; el estudio sistemático y lógico de los principios que rigen la investigación (Walker, 1985; Taylor y Bogdan, 1986).

b) Como «la lógica de los métodos». En sentido estricto, metodología significa la lógica de los métodos. Su tarea fundamental sería el estudio -la descripción,

explicación y justificación- de los métodos, y no los métodos mismos (Kaplan, 1964). En este sentido, la metodología tendría como objeto velar por los métodos, señalar sus límites y alcance, y sobre todo clarificar y valorar críticamente sus principios, procedimientos y estrategias relativas a su adecuación para la investigación (Asti Vera, 1972,22).

c) Como «proceso de investigación». La metodología es considerada también como un conjunto de operaciones conscientes, más o menos sistemáticas (fases o etapas), que se inscriben en el tiempo y cuyo rol consiste en lograr los objetivos fijados (Goyette, 1984).

El sentido que damos aquí al término *metodología* tiene un carácter general, se refiere a la manera de realizar la investigación, y más concretamente a los supuestos y principios que la rigen ..

Método. Etimológicamente el término *método* significa camino para intentar lograr un fin. Procede de la palabra griega *méthodos:* de *meta,* a lo largo, y *hodós,* camino, senda, dirección, medio, procedimiento.

Para la mayoría de los autores tiene el sentido de procedimiento o camino para conseguir un fin. Sirve de instrumento para alcanzar los fines de la investigación; su carácter regular, explícito, repetible, racional, ordenado y objetivo para lograr algo, establece el camino que la investigación ha de seguir para alcanzar su fin (Asti Vera, 1972; Bunge, 1980; Sierra Bravo, 1984).

Lo que fundamentalmente define al método es su carácter de procedimiento o conjunto de pasos sucesivos para conseguir un fin determinado. Entre sus rasgos más significativos destacan: ser una actividad sistemática, que tiene un fin y procede racionalmente.

Técnica. El término *técnica* deriva etimológicamente del griego tej*né.* Inicialmente tuvo dos sentidos: como arte práctico y como forma de actuación que se contraponía a *episteme:* la ciencia, en cuanto ésta es conocimiento o saber teórico y como tal no inmediatamente práctico. En la actualidad, el sentido de *técnica* ha quedado reservado a los procedimientos de actuación concretos y particulares, asociados a las distintas fases del método científico (Sierra Bravo, 1983,26).

Ander-Egg (1980, 32) matiza que el método es el camino, y las técnicas el arte o modo de recorrerlo. Los métodos tienen un carácter más global, abarcan varias técnicas, éstas son de carácter más práctico y operativo. Se necesitan, pues, procedimientos que hagan efectivos los métodos y esto lo hacen las técnicas. Su diferencia con el método es más de extensión que de naturaleza.

Perspectivas metodológicas

La naturaleza del problema de investigación y las preguntas relacionadas con él son los aspectos que condicionan la elección de la metodología. Existe acuerdo general en que el pluralismo metodológico es, de hecho, de importancia vital para el estudio apropiado de una cuestión de investigación. Ninguna metodología aportará por sí sola respuestas a todas las preguntas que pueden hacerse en el contexto educativo. Si

queremos saber el número de palabras que leen los niños de primero de EP, nos será de utilidad la metodología empírica; por el contrario, si queremos conocer qué significan los currícula para los alumnos universitarios, sin duda recurriremos a la metodología interpretativa.

Como la investigación aborda diferentes tipos de problemas y busca diferentes tipos de respuestas, sus procedimientos exigen diferentes metodologías. Del tipo de conocimiento que se desee alcanzar dependerá el enfoque que asuma la investigación, siendo el propósito último de ésta llegar a un conocimiento útil para la acción, ya sea política o práctica. De hecho debemos hablar de un continuo metodológico y no de polaridades opuestas.

A continuación describiremos las tres grandes perspectivas metodológicas, a modo de introducción, exponiendo los rasgos que las caracterizan. Su estudio y métodos concretos se abordarán en la segunda parte de la obra.

a) PERSPECTIVA EMPÍRICO- ANALÍTICA

Aparece también denominada con los términos *cuantitativa, positivista, científica;* toma los métodos de las ciencias fisiconaturales, considerados como el modelo del conocimiento científico. Según esta perspectiva, el objetivo de la investigación es explicar, predecir y controlar los fenómenos educativos.

Ha sido la perspectiva que ha predominado en las ciencias sociales hasta la década de los setenta, dictando los principios y criterios por los que la investigación educativa se ha regido; ha decidido y legitimado cuál es el conocimiento «verdadero», ha establecido los filtros a través de los que se legitiman los criterios de validez y fiabilidad de una investigación.

Mantiene una visión objetiva y positivista de la realidad educativa, identificándola con el mundo de los fenómenos naturales: reales, determinados, externos al sujeto. Predomina el proceso hipotético-deductivo y se busca la generalización de los resultados a partir de muestras de población representativas.

Desde esta perspectiva la investigación educativa se propone el estudio de relaciones y regularidades con el fin de descubrir las leyes universales que explican y rigen la realidad educativa.

La metodología empírico-analítica participa de los supuestos del positivismo y de la ciencia nomotética que tiende a centrarse en las manifestaciones «externas» de la realidad educativa considerada hasta cierto punto repetitiva, predecible e invariable. En general, se reduce a los fenómenos observables que sean susceptibles de medición, análisis estadístico y control experimental.

La metodología empírico-analítica hace referencia al conjunto de métodos cuya lógica de justificación se apoya en los principios de objetividad, enfatizando la evidencia empírica y cuantificación.

La «objetividad» es un rasgo básico de esta perspectiva; se refiere al proceso metodológico de la investigación. Se entiende «como acuerdo entre jueces competentes» (personas, instrumentos o mecanismos). La objetividad hace referencia a que todos los procedimientos de la investigación deben ser públicos, replicables e independientes del investigador. Afecta a todo el proceso de investigación, desde la planificación y recogida de datos hasta el análisis e interpretación de éstos.

La «evidencia empírica» supone que la investigación debe estar guiada por la evidencia obtenida directa o indirectamente de la observación. Se constituye en el criterio de contraste o falsación de hipótesis y enunciados científicos. La evidencia empírica sustituye a las fuentes de autoridad, sentido común o razonamiento especulativo.

En la metodología empírica la «cuantificación» es una consecuencia directa de la medida, que por definición produce datos cuantitativos cuyo análisis exige la utilización de modelos estadísticos. Pretende generalizar los resultados, por supuesto a nivel probabilístico, por lo que acude al análisis estadístico (De la Orden, 1989,225-226).

En las últimas décadas, desde diferentes sectores educativos -investigadores, académicos, prácticos, administrativos, etc.-, se vienen produciendo continuas críticas que cuestionan la aplicabilidad y utilidad de esta metodología para las ciencias de la educación, dada la complejidad de la realidad social y las exigencias de rigor y control que exigen los diseños experimentales (De Miguel, 1988).

Esta metodología es considerada insuficiente para explicar la complejidad de la realidad educativa comparada con el orden y regularidad del mundo natural. Col1tempta_ la educación de forma mecanicista y reduccionista, ignorando aspectos tan significativos como son la capacidad para elegir, la libertad, la individualidad y la responsabilidad (Keeves, 1988). Además, ciertos ámbitos educativos son difíciles de observar sin distorsionarlos; muchas situaciones educativas son irrepetibles y sus resultados son de escasa utilidad para la praxis educativa por su artificialidad.

Los párrafos siguientes son un resumen de los principales rasgos que definirían la naturaleza de la metodología estrictamente empírico-analítica.

- *Naturaleza de la realidad.* Considera la realidad como algo externo al investigador, singular y tangible, que puede fragmentarse en variables.
- *Finalidad de la investigación.* Su cometido es conocer y explicar la realidad para predecirla y controlarla. Pretende llegar a generalizaciones libres del tiempo y contexto, con el fin de formular las leyes y explicaciones nomotéticas que rigen los fenómenos naturales.
- *Naturaleza de la relación investigador-objeto investigado.* El investigador es visto como un ser objetivo, apolítico, libre de valores, que trabaja distanciado del objeto de estudio.
- *Problemas que investiga.* La mayoría de las veces los problemas surgen de las teorías o postulados. Se orienta a contrastar teorías. Tiene una orientación eminentemente teórica, contempla menos las necesidades de los individuos.
- *Papel de los valores.* La investigación empírico-analítica mantiene que la metodología está libre de valores. El método es garantía de neutralidad.
- *Teoría y práctica.* Separa la teoría de la práctica. La teoría tiene un carácter de normativa para la práctica; ésta queda supeditada a los cánones que dicta la teoría.
- *Criterios de calidad.* Establece como criterios de calidad la validez, la fiabilidad y la objetividad.
- *Instrumentos.* Se basa en instrumentos que implican la codificación (cuantificación) de los hechos, utilizando tests, cuestionarios, escalas de medida, entrevistas estructuradas, etc.
- *Análisis de los datos.* Es de carácter deductivo y estadístico. Aporta análisis cuantitativos.

b) PERSPECTIVA HUMANÍSTICO- INTERPRETA TIV A

Vinculada a la perspectiva antipositivista se encuentra la metodología interpretativo-humanista o cualitativa (Marshall y Rossman, 1989), que se constituye como una metodología alternativa al enfoque empírico-analítico anterior.

Los investigadores que se inclinan por este enfoque perciben la esfera educativa mucho más flexible y personal, creada por los propios sujetos. La realidad sólo puede estudiarse recurriendo a los puntos de vista de los sujetos implicados en las situaciones educativas. En contraposición al observador externo, objetivo e independiente, ahora, la realidad educativa puede ser analizada con más propiedad por el investigador que comparte el mismo marco de referencia que las personas investigadas. La comprensión y valoración de las interpretaciones del individuo de la realidad y de las situaciones educativas en las que se halla implicado han de venir del propio sujeto, no del exterior.

Para esta perspectiva, el proceso educativo tiene un carácter subjetivo y por ello es necesaria la experiencia directa de las personas en contextos educativos específicos. Las personas se mueven por sus creencias e interpretaciones del medio (Jorgensen, 1989). Se estudian aspectos de la educación como la motivación, intenciones, acciones y significados, difícilmente observables y cuantificables.

Este enfoque se interesa por los significados e intenciones de las acciones humanas.

Se centra en las personas y analiza las interpretaciones que hace del mundo que le rodea y de su relación con él (Denzin, 1989). Su objetivo es lograr imágenes multifacéticas del fenómeno que hay que estudiar tal como se manifiesta en las distintas situaciones educativas y contextos implicados. Desde esta perspectiva se contempla el mundo subjetivo de la experiencia humana. El investigador intenta penetrar en el interior de la persona y entenderla «desde dentro» (Marshall y Roosman, 1989).

En consecuencia, el investigador sigue un enfoque holístico-inductivo- idiográfico, se interesa por la comprensión global de las situaciones y personas; utiliza la vía inductiva; los conceptos, comprensiones e interpretaciones se elaboran a partir de los datos; el investigador crea un marco de referencia adecuado para que las personas puedan responder fielmente sobre el mundo según sus experiencias y vivencias; tiene en cuenta la idiosincrasia de los fenómenos y contextualiza las situaciones.

Utiliza diseños de naturaleza flexible que llevan a un enfoque progresivo; la metodología está al servicio del investigador y no a la inversa. Tiende a utilizar técnicas cualitativas, como la observación participante, entrevistas, notas de campo, relatos. Reinterpreta los conceptos de validez, fiabilidad y objetividad propios de la metodología empírico-analítica. Se apoya en los procesos de triangulación y contratación intersubjetiva.

La finalidad de la investigación será comprender cómo los sujetos experimentan, perciben, crean, modifican e interpretan la realidad educativa en la que se hallan inmersos. Observando a las personas en su contexto natural y diario, entrevistándolas y analizando sus relatos y documentos, se obtiene un conocimiento directo de la realidad educativa, no filtrado por definiciones conceptuales, operativas y escalas previamente estructuradas.

En los párrafos siguientes se resumen los principales rasgos que definen la naturaleza de la metodología interpretativa.

- *Naturaleza de la realidad.* Para la corriente interpretativa la realidad se presenta múltiple e intangible, que sólo se puede abordar de forma holística; es algo que se construye.

- *Finalidad de la investigación.* Comprender e interpretar los significados de los fenómenos y acciones sociales.
- *Naturaleza de la relación investigador-objeto investigado.* La relación entre investigador y sujetos investigados se admite, se da interrelación, el investigador suele participar y actuar. - *Problemas que investiga.* En este tipo de investigación los problemas están relacionados con las necesidades del grupo social. Su propósito es comprender la situación estudiada desde el punto de vista de los sujetos.
- *Papel de los valores.* Admite la influencia de los valores en la investigación. El investigador no está libre de valores, sino que los explicita.
- *Teoría y práctica.* Se da un intercambio dinámico entre teoría y práctica, con retroinformación y modificaciones constantes de la teoría en base a los datos obtenidos. Se establece un proceso dialéctico entre ambas.
- *Criterios de calidad.* Como criterios de calidad propone: credibilidad, transferencia, dependencia y confirmabilidad.
- *Instrumentos.* Utiliza estrategias de naturaleza cualitativa, como la observación participante, la entrevista informal, los diarios, los registros de campo, el análisis de documentos, etc.
- *Análisis de los datos.* Es de naturaleza cualitativa; suele implicar varias etapas: reducción, categorización, representación, validación e interpretación. Es un proceso de carácter cíclico interactivo.

Las críticas que se vierten sobre esta metodología suelen referirse a su carácter subjetivo o riesgo de subjetividad; las observaciones e informes de los sujetos pueden estar sesgados o ser incompletos. El mismo investigador puede analizarlas situaciones desde su propio marco referencial.

Suele descuidar la fiabilidad, haciendo más hincapié en la validez. Los datos suelen carecer de consistencia por falta de técnicas de replicabilidad. La debilidad de los datos es algo inherente a la metodología interpretativa.

c) PERSPECTIVA ORIENTADA A LA PRÁCTICA EDUCATIVA: TOMA DE DECISIONES Y CAMBIO

A mediados de los sesenta surge en el campo de la investigación educativa un importante y creciente interés por los estudios relacionados con la práctica educativa, dando origen a una perspectiva de investigación denominada *investigación orientada a la política (policy-oriented research),* en contraposición a la investigación fundamental orientada a ampliar el cuerpo de conocimiento existente (Nisbet, 1988).

La investigación orientada a la práctica educativa se define mejor en términos de su función instrumental que por los tópicos que estudia. Con esta expresión nos referimos a la investigación educativa que se diseña, realiza y comunica con el propósito específico de proporcionar información para la toma de decisiones (a nivel político o dentro de la práctica escolar), de controlar la implantación de una determinada política o de examinar los efectos de la política existente (Keeves, 1988; Nisbet, 1988).

En su sentido estricto se refiere a la investigación que tiene una aplicación directa en la política o práctica educativas; en sentido amplió se extiende a los estudios de investigación diseñados con el propósito de comprender los procesos educativos y mejorar la praxis educativa (Nisbet, 1988).

La nota esencial de esta perspectiva es que la investigación se diseña para contribuir a solucionar los problemas o aportar directrices para la acción, describiendo lo más ampliamente posible la complejidad de las situaciones y estableciendo marcos conceptuales que posibiliten la toma de decisiones con el mayor *insighty* comprensión posibles. De modo que el resultado final del estudio son las decisiones y recomendaciones para la acción, y no tanto su contribución a la creación de conocimiento o teoría.

Así pues, lo que distingue a la investigación orientada a la política educativa de otras perspectivas es el propósito para el que es diseñada, y no el tema elegido o la metodología aplicada. Queremos recalcar que esta perspectiva no posee una metodología propia en el sentido de las dos perspectivas anteriores, sino que se sirve de sus métodos, lo que no quiere decir que no posea sus propios diseños de investigación.

Frente a la perspectiva empírico-analítica centrada en la explicación de las conductas, o la humanístico-interpretativa en la comprensión de las acciones, la investigación orientada a la política educativa tiene como fin optimizar la práctica educativa mediante la adquisición de conocimiento práctico. Si bien normalmente opera dentro de un contexto o teoría adoptada, sin embargo, su propósito no es tanto cambiar la teoría o generalizar los resultados como impactar sobre la política o práctica escolar para la que fue diseñada.

Dentro de esta perspectiva se agrupan diversas modalidades de investigación, desarrolladas con la finalidad de aportar información que guíe los procesos de cambio y toma de decisiones con el propósito de introducir mejoras en el campo del trabajo social en general y en la educación en particular. Así, la investigación participativa *(participatory research)* pone el acento en el cambio social y desarrollo comunitario, mientras que la investigación legitimadora *(legitimatory research)* se interesa por evitar el cambio u obtener evidencia como soporte del sistema o política educativa existente. La investigación acción *(action research)* es conducida por los propios prácticos con la intención de estudiar y reflexionar sobre *su propia práctica educativa.* En el campo de la investigación evaluativa el interés por mantener el sistema o propiciar el cambio depende del contexto social o político en el que se realiza la evaluación, que a su vez influirá en el modelo de evaluación elegido.

La perspectiva de investigación orientada a la política educativa incluye toda clase de estudios que aporten datos o evidencia empírica que orienten la toma de decisiones informada por los resultados de la investigación y no por los prejuicios o suposiciones. Incluye los estudios orientados a indagar las soluciones que presentan los acuciantes problemas sociales o educativos, identificando y dando respuesta a los problemas implicados en la implantación de decisiones políticas, el control y evaluación de las innovaciones educativas, los estudios experimentales o la comparación de métodos educativos alternativos.

El análisis de las políticas educativas ha ido creciendo en los últimos años en el campo de la investigación educativa, abarcando estudios tales como:

a) estudios de muestras para recoger «hechos» relevantes como base de datos para la decisión;
b) estudios experimentales para resolver controversias;
c) estudios de desarrollo para implementar políticas educativas;
d) estudios de evaluación (Nisbet, 1988, 141).

2
2
2
2
2

PARTE II

METODOLOGÍAS DE INVESTIGACIÓN EDUCATIVA DE PERSPECTIVA EMPÍRICO-ANALÍTICA

Recordemos que la perspectiva empírico-analítica deriva de los enfoques empleados en las ciencias físiconaturales. Tiende a centrarse más en aspectos cuantificables de los fenómenos educativos con el fin de constatar relaciones y explicaciones causales generalizables, es decir, enfatiza más el contexto de justificación o contrastación de hipótesis. Los problemas planteados en la orientación empírico-analítica suelen requerir datos cuantitativos, obtenidos con instrumentos estructurado s válidos y fiables, y un análisis de datos con predominio de procedimientos matemáticos o estadísticos. Es necesaria la replicabilidad en los datos recogidos y en el análisis realizado, destacándose la naturaleza nomotética de la investigación.

Aunque esta perspectiva metodológica no pueda abordar los múltiples aspectos de la realidad educativa, sus aportaciones son tan valiosas que sigue siendo la orientación predominante en determinados ámbitos y situaciones educativas.

En primer lugar, expondremos algunas consideraciones generales para enmarcar distintas metodologías de orientación empírico analítica. Posteriormente, y en el contexto de cada metodología, pasaremos a describir diferentes diseños y estrategias de investigación.

CAPÍTULO 4

Consideraciones generales

Antes de abordar la diversidad metodológica que caracteriza esta perspectiva nos centraremos en algunos conceptos fundamentales como las características que debe tener el diseño, la noción de control y la selección de la vía metodológica más adecuada para aportar información sobre el problema planteado.

4.1 CARACTERÍSTICAS DEL DISEÑO

El diseño estructura la organización de la investigación y es un esquema global que indica lo que realizará el investigador, cómo alcanzará los objetivos de la investigación y cómo abordará el problema planteado (Kerlinger, 1975, 214, Y Bartolomé, 1988, 102).

Como ejemplo ilustrativo, supongamos que **un** educador quiere saber si **un** grupo de niños que aprenden a leer con el método sintético (al) logran mayor eficiencia lectora que los sometidos al método analítico (a). Para ello podemos estudiar la relación entre el método de enseñanza (VI) y el aprendizaje de la lectura (VD) a través de **un** diseño como el de la tabla 4.1.

TABLA 4.1 Representación simbólica de un ejemplo de diseño para contrastar la influencia de dos métodos de lectura

Grupos	*n*	*Asignación*	*Pretest*	*Método (VI)*	*Postest (VD)*
1	n_1	No aleatoria	—	a_1	$\bar{X}_1$
2	n_2	No aleatoria	—	a_2	$\bar{X}_2$

El diseño constituye la organización formal básica de la investigación (Pereda, 1987, 250) Y en su representación simbólica (tabla 4.1) puede incluirse cierta información sobre los siguientes aspectos:

a) Variables implicadas en el problema y su operativización. La variable independiente presenta dos modalidades de enseñanza basadas respectivamente en los métodos sintético y analítico. Además, el diseño sugiere que el aprendizaje de la lectura se medirá cuantificando el nivel lector adquirido por los sujetos, probablemente a través de algún instrumento estructurado.

b) Grupos y número de sujetos de cada grupo. Los distintos grupos se denominarán con cifras correlativas: 1, 2, 3, etc. En la columna encabezada por el símbolo *n* se indicará el número de sujetos que componen el grupo respectivo.

c) *Asignación de los sujetos a los grupos.* En el diseño presentado no se han asignado al azar a cada grupo, es decir, son grupos que pueden diferir entre sí en otras variables. En los diseños con más de un grupo se indica si los grupos se han formado aleatoriamente o se recurre a grupos naturales o estáticos que no han podido aleatorizarse.

d) Especialmente cuando en el diseño intervienen más de una variable independiente conviene simbolizar las *categorías o niveles* de dichas variables. Con letras latinas mayúsculas se indica una variable o factor y con minúsculas las distintas categorías o niveles con sus respectivos subíndices. Por ejemplo:

Variables independientes	*Categorías* o *niveles de las variables independientes*		
A = Método de lectura	a_1 = Sintético,	a_2 = Analítico	
B = Tiempo de estudio	b_1 = 1 hora,	b_2 = 2 horas,	b_3 = 3 horas

e) Fase de la medición de la variable dependiente. Antes (pretest) y después de intervenir la variable independiente (postest). Las distintas medidas se expresan del siguiente modo:

$$X_1,\ X_2, X_3, X_4, \text{etc.}$$

En el diseño de la tabla 4.1 no se ha medido el nivel lector antes de aplicar los métodos de lectura.

f) Análisis estadístico. Habrá que realizar, en este caso, una comparación de las medias obtenidas por cada grupo en lectura.

El diseño más adecuado para una investigación debe poseer los requisitos de validez, fiabilidad, simplicidad y grado de significación (Cox, 1958; Rosel, 1986, y Pereda, 1987).

Validez

Un diseño tiene validez cundo permite detectar la relación real que pretendemos analizar. Pueden contemplarse tres tipos de validez.

a) Validez interna. El diseño posee validez interna si existen garantías de que la relación encontrada entre las variables estudiadas no se debe a la presencia de otras variables. Por ejemplo, al finalizar un curso de eficiencia lectora los alumnos de primero A de secundaria obtienen un nivel lector estadísticamente superior al de primero B. Si la clase de primero A tiene un nivel medio de inteligencia estadísticamente superior al de primero B, el investigador no podrá concluir que el curso de eficiencia lectora aumenta significativamente el nivel lector, ya que la diferencia encontrada puede ser debida a que el grupo de primero A tiene más inteligencia que el grupo de primero B. En consecuencia, el diseño carece de validez interna.

b) Validez externa. Se refiere a la representatividad o generalización de los resultados de una investigación, es decir, a la posibilidad de generalizar los resultados a otros sujetos, a otros grupos, a otros tratamientos y condiciones, a otras variables dependientes y a otras situaciones con textuales reales. Este tipo de validez es necesario porque es imposible recoger todos los datos posibles. Al generalizar pueden contemplarse, al menos, tres posibilidades:

Validez de población: Posibilidad de hacer extensibles a la población los resultados obtenidos con la muestra elegida.

Validez contextual o ecológica. Referida a la posibilidad de que los resultados obtenidos en la situación de investigación sean aplicables a situaciones educativas reales.

Validez de los tratamientos. Cuando las categorías de la variable independiente son representativos de los niveles que adopta dicha variable en la realidad. Así, cuatro, cinco y seis horas seguidas de lectura en niños de ocho años serían valores poco representativos de la capacidad de lectura real en estas edades.

c) *Validez conceptual.* Las definiciones operativas de las variables implicadas deben ser coherentes con las definiciones conceptuales de las mismas, tal como son delimitadas por la teoría de la que proceden. Así, cuando conceptualizamos la inteligencia como capacidad verbal debe operativizarse o medirse con una prueba verbal.

Fiabilidad

La fiabilidad o precisión del diseño se refiere a la constancia para captar la relación entre las variables. En consecuencia, la fiabilidad del diseño se favorece si se eligen valores adecuados en las variables, se aplican convenientemente y se miden con precisión. La fiabilidad de un diseño suele estimarse al aplicar sucesivas veces el mismo diseño en circunstancias lo más similares posible. A partir de estas replicaciones puede verse si se obtienen los mismos resultados.

Simplicidad

El diseño no se debe complicar más allá de lo necesario para aportar información sobre el problema y obtener una conclusión inequívoca de la investigación. Así, se tomarán los datos necesarios, procedentes de una muestra suficientemente grande y con el menor número posible de grupos.

El investigador procurará que el problema que ha de investigar se defina con precisión y concreción, facilitando la búsqueda o elaboración de un diseño coherente con el problema que ha de investigar y la conclusión que se pueda sacar del mismo será menos ambigua.

Nivel de incertidumbre

Para poder comprobar o descubrir una relación entre fenómenos debemos operativizar las variables implicadas y asignar un nivel de incertidumbre de las conclusiones por medio de la hipótesis de nulidad estadística. La hipótesis de la investigación se acepta o rechaza en función del nivel de significación adoptado.

4.2 CONTROL DE LAS VARIABLES

En la perspectiva empírico-analítica los problemas de investigación suelen requerir que el investigador trate de constatar relaciones y explicaciones de los fenómenos educativos. Esto exige cierto grado de control o manipulación de las variables implicadas con el fin de garantizar la validez interna de la investigación. El control se refiere, por tanto, al poder o dominio que tiene el investigador sobre las condiciones en que se desarrolla su investigación. Para ejercer el control se utilizan distintos procedimientos que permiten clarificar si la variable independiente ejerce alguna influencia en la variable dependiente.

El control se ejerce a través de tres posibles actuaciones del investigador:

1) *Control de la variable independiente.* Puede ejercerse de dos maneras: manipulando o seleccionando los valores de la variable independiente.

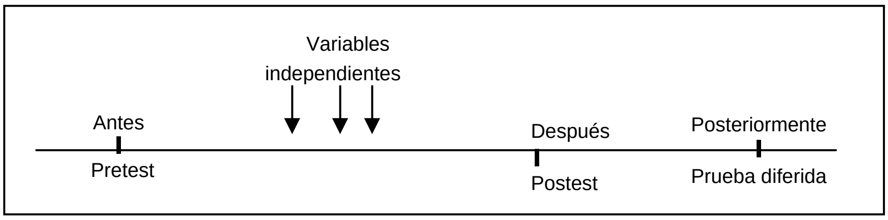
a) *Manipulación.* Constituye el máximo grado de control y consiste en variar deliberadamente valores de la variable independiente para ver qué cambios aparecen en la variable dependiente. El investigador determina los valores de la variable independiente y qué grupos de sujetos recibirán dichos valores, es decir, interviene provocando la situación de investigación. Así, puede determinar las condiciones y formas de comportarse un profesor en clase; por ejemplo, puede alabar ciertos comportamientos y recriminar otros, gratificar o castigar. También puede utilizar audiovisuales o recursos informáticos para favorecer el aprendizaje. De la misma manera, el investigador puede someter a tres grupos de sujetos, respectivamente, a 5, lO y 20 minutos de ejercicios de psicomotricidad.

b) *Selección.* Un grado menor de control se ejerce cuando la variable independiente no es susceptible de manipulación, o no es viable su manipulación. En estos casos, el investigador elige los sujetos que ya poseen los valores deseados de la variable independiente. Así, con variables como la clase social, el investigador ha de limitarse a seleccionar niveles como pueden ser alta, media y baja.

2) Control de la variable dependiente. Para controlar la variable dependiente debe medirse dicha variable. Hay que determinar el número de medidas y el momento en que se lleva a cabo la medición.

a) *Número de medidas.* Puede efectuarse una sola medida de la variable dependiente, dos o más medidas.

b) *Momento de la medición.* Si se toma como punto de referencia el período de aplicación de la variable independiente, puede medirse la variable dependiente antes (pretest), inmediatamente después (postest) y, cuando quiere determinarse la permanencia del efecto, posteriormente (prueba diferida).



3) Control de las variables extrañas. Durante el experimento pueden presentarse una serie de influencias procedentes de variables extrañas que, junto a la variable independiente, pueden llegar a influir de manera decisiva en la variable dependiente. El control de las variables extrañas consiste en eliminar, o bien igualar, la posible incidencia de variables ajenas a los objetivos de la investigación. Por ejemplo, el investigador procura conseguir la equivalencia entre los grupos que van a ser sometidos a distintos niveles de la variable independiente para que luego sean comparables los efectos. Los cambios observados en la variable dependiente sólo son atribuibles a los cambios introducidos en la variable independiente cuando las variables extrañas están debidamente controladas, es decir, cuando influyen por igual en cada grupo sometido a un nivel de la variable independiente. Esto se consigue cuando se han tomado las debidas precauciones para que dichos cambios no sean producidos por otras variables extrañas, procedentes de la situación, como el ruido y la temperatura, o de los sujetos, como la edad, inteligencia y motivación.

Para clarificar mejor la importancia del control volvamos al ejemplo presentado en la tabla 4 .1. Queríamos comprobar si los niños que aprendieron con él método sintético (grupo 1) logran mayor eficiencia lectora que los sometidos al método analítico (grupo 2). Para sacar una conclusión sobre la relación entre las variables método y nivel lector hay que controlar los efectos de variables ajenas al propósito del estudio, es decir, el investigador ha de eliminar o igualar los efectos diferenciales de variables extrañas que puedan afectar al nivel lector. Así, la inteligencia podría crear un efecto diferencial, ya que pudiera favorecer más el nivel lector de un grupo que el de otro;<en consecuencia, debería ser considerada como una variable extraña que habría que controlar. Supongamos que el grupo 1 tuviera un promedio de inteligencia estadísticamente superior al del grupo 2 (tabla 4.2).

La posible diferencia entre las medidas de lectura no puede atribuirse sólo a la diferencia entre la eficacia de los métodos. Es decir, la variación (varianza) observada en el nivel lector (VD) podría deberse a la variación que hemos introducido en los métodos (VI), pero también a la que ya había en la inteligencia de los grupos. Hay una fuente de variación extraña que distorsiona los efectos y nos impide que atribuyamos la diferencia en lectura sólo al método empleado. En este caso el investigador no puede

separar lo que se debe al método de lo que se debe a la inteligencia. Si obtuviera diferencias significativas entre los niveles medios de lectura no sabría si la relación de causalidad puede establecerse:

1) Entre el tipo de método y el nivel lector.
2) Entre la inteligencia y el nivel lector.
3) Una combinación de ambas posibilidades.

TABLA 4.2 Ejemplo de diseño para contrastar la eficacia de dos método de lectura ni controlar la inteligencia

Grupos	*Promedio inteligencia*	*VI*	*VD*
1	95	Sintético	$\bar{X}1$ en lectura
2	90	Analítico	$\bar{X}_2$ en lectura

La confusión desaparece al controlar el efecto de la variable extraña. El investigador controla la inteligencia cuando vuelve comparables los grupos 1 y 2 con respecto a dicha variable. Para ello puede formar los grupos 1 y 2 con sujetos que tengan la misma inteligencia o puede distribuir los sujetos al azar, de manera que unos con otros compensen las diferencias. Al actuar así obtiene dos grupos equivalentes en inteligencia (tabla 4.3) y en las demás variables.

TABLA 4.3 Diseño para contrastar la eficacia de dos métodos de lectura controlando la inteligencia

Grupos	*Promedio inteligencia*	*VI*	*VD*
1	93	Sintético	$\bar{X}1$ en lectura
2	93	Analítico	$\bar{X}_2$ en lectura

En consecuencia, el investigador debe controlar diferencias (variaciones) introducidas previamente por los sujetos (inteligencia, motivación, edad) o procedentes de la situación (hora, ruido, luminosidad). Sólo así tendrá la seguridad de que las diferencias observadas en la variable dependiente serán atribuibles a los niveles de la variable independiente.

Las técnicas de controlo procedimientos para aumentar la equivalencia entre los grupos se abordarán más adelante al tratar los diseños experimentales.

4.3 METODOLOGÍAS EMPÍRICO-ANALÍTICAS

Las modalidades de investigación educativa, derivadas del método científico (bajo el paradigma positivista, o también denominado racionalista o normativo), tienden a

utilizar la metodología nomotética, basada en la perspectiva cuantitativa, y pueden agruparse en tomo a tres grandes modalidades que constituyen niveles distintos de control y, por ello, de aproximación a la explicación: metodología experimental, cuasiexperimental y no experimental o ex-post-facto.

Metodología experimental	Metodología cuasiexperimental	Metodología no experimental

Partimos, por tanto, de la noción de experimentación entendida como provocación intencional de la situación que hay que investigar, con la posibilidad de utilizar distintos grados de manipulación y control de las variables extrañas. En la tabla 4.4 pueden apreciarse las tres modalidades de investigación según el grado de control ejercido.

TABLA 4.4 Metodologías de investigación según el grado de control ejercido

Enfoque	*Grado de control*
Experimental	ALTO. Se provoca (manipula) el fenómeno, el investigador de termina los valores de VI según su conveniencia. Existe un control máximo de todas las variables extrañas más significativas.
Cuasiexperimental	MEDIO. Se provoca o manipula el fenómeno, el investigador determina los valores de VI según su conveniencia. Quedan por controlar muchas variables extrañas significativas.
No experimental Ex-post-facto	BAJO. Actitud pasiva. No se modifica el fenómeno o situación objeto de análisis, pues la relación entre las variables ya se ha producido con anterioridad y el investigador sólo puede registrar sus medidas.

4.4 CRITERIOS DE SELECCIÓN DE LA METODOLOGÍA

De acuerdo con lo expuesto pueden ofrecerse algunas directrices generales sobre la conveniencia de una vía metodológica para aportar información sobre un problema de investigación. Algunos criterios orientativos son los siguientes:

1) *Grado de control de la variable independiente y de las variables extrañas.* No existe una clara delimitación entre las tres modalidades de investigación. Puede

hablarse de un verdadero continuo metodológico (tabla 4.5) en cuanto a dos aspectos básicos que determinan grados de control progresivamente menores de izquierda a derecha.

TABLA 4.5 Continuo metodológico según el grado de control

Enfoque experimental	Enfoque cuasiexperimental	Enfoque no experimental
Control óptimo Provocación y manipulación del fenómeno	Control medio Se manipula VI, pero no se controlan otras variables extrañas significativas	Control bajo El investigador permanece pasivo, sólo selecciona

El enfoque no experimental o ex-post-facto es más adecuado cuando el investigador sólo puede seleccionar y decidir qué características se han de observar o medir. En cambio, si puede provocar el fenómeno y controlar las variables extrañas, a través del experimento podrá abordar mejor las relaciones causa efecto.

2) Validez interna y externa. De acuerdo con lo expuesto, podemos afirmar que a medida que nos aproximamos al ideal del experimento estricto tenderá a aumentar la validez interna y el control, mientras que la validez externa disminuirá debido a la artificialidad de la situación. El proceso se invertirá cuando el investigador se vea obligado a orientarse hacia un estudio no experimental o ex-post-facto.

3) Naturaleza de la situación de investigación. Cuando interesa realizar la investigación en un contexto real, para aumentar la validez externa, aunque manteniendo un grado razonable de validez interna, es preferible optar por la metodología cuasiexperimental o incluso ex-post-facto. Las posibles deficiencias en el control de las variables extrañas que conlleva la situación real puede quedar compensada por la mayor proximidad a la realidad. En cambio, cuando el interés se centra más en la pura relación de causalidad entre variables, en la contrastación de una hipótesis, es aconsejable una situación de laboratorio, cuya artificialidad posibilita el máximo grado de control, tan característico de los experimentos estrictos.

4) Objetivo o propósito del investigador. Para aportar información sobre los problemas planteados, y teniendo en cuenta la naturaleza de los mismos, el investigador puede formular objetivos como los siguientes:

a) *Describir relaciones entre los fenómenos.* La descripción consiste en detectar los componentes de los fenómenos, sus características, su nivel de aparición y el grado de variación conjunta que puedan presentar. En general, este objetivo es característico de las metodologías que hemos englobado bajo la denominación de ex -post-facto. Para satisfacer este objetivo el investigador puede realizar acciones .como las siguientes:

- *Comparar:* Recoger información a través de la contrastación de más de un conjunto de datos. Se efectúan juicios sobre si existen o no diferencias, sobre si un conjunto de datos refleja un mayor nivel que otro en la posesión de alguna

característica o variable. Por ejemplo, pueden efectuarse comparaciones entre proporciones y entre medias.

- *Asociar:* Constatar el grado de variación conjunta que presentan dos o más conjuntos de datos nominales u ordinales. Es decir, grado en que al aparecer las categorías o rangos de una variable tienden a presentarse también en las categorías o rangos de la otra. Algunos ejemplos de coeficientes para estimar este tipo de relación pueden ser: ji-cuadrado, contingencia y phi.
- *Correlacionar:* Constatar el grado de variación conjunta que presentan dos o más variables cuantitativas. Es decir, grado en que al variar los valores de una variable tienden también a variar los de la otra.

b) *Predecir categorías o valores de fenómenos.* Predecir es indicar en qué condiciones ocurren los fenómenos, con cierto grado de probabilidad. Cuando a partir de los valores de unas variables pueden estimarse los valores de otra se utiliza un método predictivo.

c) *Explicar relaciones de causalidad entre fenómenos.* Establecer relaciones entre los fenómenos, determinando su dirección y sentido. Detectar qué variables provocan cambios en los valores de otra variable. Para satisfacer este objetivo puede ser necesario:

- *Comparar.* Cuando es viable provocar el fenómeno intencionalmente y es posible formar los grupos aleatoriamente, el investigador puede aplicar el método experimental y comparar datos procedentes de grupos sometidos a la variable independiente. Para estimar la posible relación de causalidad, el investigador ha de controlar la independiente, es decir, ha de modificar la producción del fenómeno o situación, manipulando ciertas variables con el fin de comparar los efectos que se producen. Para ello ha de eliminar o neutralizar (equilibrar) los efectos de otras variables que junto a la variable independiente pueden ejercer influencia diferenciada y sus efectos podrían atribuirse equivocadamente a la variable independiente.

  Si los datos que hay que comparar proceden de una situación no susceptible de aleatorización, el investigador actuará a un nivel cuasiexperimental, y cuando además no sea viable provocar intencionalmente el fenómeno, el método se denomina *comparativo-causal*, ya que el investigador compara con intención de explorar relaciones de causalidad, siempre con la precaución que exige un nivel ex-post-facto.
- *Correlacionar.* Aunque la existencia de correlación entre variables es una condición necesaria pero no suficiente para explorar relaciones de causalidad, el investigador puede recurrir a los denominados *modelos estructurales causales* con el fin de realizar inferencias causales.

Otros objetivos, como comprender un fenómeno desde la perspectiva de las personas implicadas, valorar la implantación y los efectos de una intervención u optimizar un proceso, serán contemplados respectivamente en las perspectivas humanístico-interpretativas y en la orientada a la toma de decisiones y al cambio.

Para tomar una decisión sobre qué método es más adecuado ante un problema planteado hay que conjugar los criterios mencionados y optar por la vía metodológica

que responda mejor a las necesidades de control, a los objetivos del investigador, a las exigencias de la validez y a la naturaleza de la situación en la que se produce el fenómeno. En la tabla 4.6. puede verse una síntesis de algunos de los criterios mencionados en el contexto de la orientación empírico-analítica.

En consecuencia, conviene tomar una opción metodológica que conjugue de manera óptima los intereses y posibilidades del investigador con las exigencias del problema planteado y la naturaleza de la situación de investigación.

TABLA 4.6 Algunos criterios orientativos para seleccionar la metodología más adecuada

Metodología	*Grado de validez interna*	*Grado de validez externa*	*control*	*Situación de investigación*	*Orientación temporal*
Experimental	Mayor	Menor	Mayor	Artificial	Futuro
Cuasiexperimental	Media	Mayor	Medio	Natural	Futuro
Ex-post-facto	Menor	Mayor	Menor	Natural	Presente

En resumen, de lo comentado hasta aquí se desprenden las siguientes consideraciones y criterios referenciales:

1) El primer gran núcleo del continuo metodológico está constituido por el método estrictamente experimental, donde el investigador introduce deliberadamente una variación o manipulación en al menos una variable independiente, para ver qué efectos produce en la variable dependiente, en condiciones rigurosamente controladas con el fin de evitar que dichos efectos queden contaminados por la influencia diferenciada de otras variables. El método experimental sigue el modelo hipotético-deductivo de la ciencia y posibilita la mejor aproximación a las relaciones causa-efecto. Si se tiene en cuenta que la ciencia pretende establecer leyes que evidencien relaciones causales entre variables, los métodos experimentales serán los más adecuados para explicitar de forma precisa y controlada dichas relaciones causa-efecto entre los fenómenos, siempre que la variable independiente sea susceptible de modificación intencional por parte del investigador.

2) En un ámbito menos estricto, los diseños cuasiexperimentales han tenido difusión también en áreas como la sociología de la educación y la pedagogía diferencial. Este núcleo intermedio denominado *métodos cuasiexperimentales* se refiere a situaciones de investigación provocadas por el investigador, pero sin tener un control estricto de las variables extrañas a los propósitos de la investigación.

3) En el otro extremo del continuo habíamos situado los métodos no experimentales o ex-post-facto, así denominados porque no implican una modificación o

producción del fenómeno que hay que investigar, es decir, en sentido amplio, se limitan a describir una situación que ya viene dada al investigador, aunque éste pueda seleccionar valores para estimar relaciones entre las variables. La obtención de los datos se lleva a cabo después de que la relación entre las variables implicadas ya se haya producido en su contexto natural y, por tanto, el investigador no puede modificar ningún valor de la variable independiente. Esta metodología se orienta más a una situación actual que puede reflejar la relación entre variables cuya influencia se produzca en el momento presente o se haya producido en el pasado.

En educación es frecuente que algunos métodos de investigación tengan que situarse en esta posición más próxima al otro extremo del continuo metodológico mencionado. Así, puede ocurrir que el investigador esté interesado en explorar relaciones de causalidad pero no pueda provocar el fenómeno, bien porque las variables independientes no son susceptibles de manipulación, bien porque la situación de investigación ya ocurrió, o bien porque siendo manipulables no resulta viable provocar la situación de investigación por razones personales o contextuales. Ahora bien, aunque no se provoque el fenómeno que hay que investigar, el investigador puede seleccionar o tomar directamente grupos estáticos o naturales y posteriormente llevar a cabo comparaciones.

CAPÍTULO 5

Metodología experimental

El proceso general de la investigación ya se ha descrito en el capítulo 3; sin embargo, cada método presenta variaciones que le permiten adaptarse a las exigencias de problemas específicos de investigación. En consecuencia, es apropiado apuntar algunas características inherentes al método experimental, abordando posteriormente aspectos como la varianza y las técnicas para controlarla, las fuentes de invalidez y la estrategia que hay que seguir en los diseños más importantes. A su vez, se contemplan las posibilidades y límites de dichos diseños y se ofrecen orientaciones para planificar un experimento.

Los estudios experimentales estrictos se caracterizan por presentar los rasgos siguientes:

1) El investigador manipula la variable independiente y le asigna niveles.

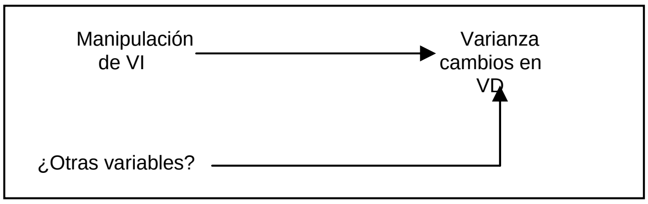
2) Han de aplicarse dos o más niveles de la variable independiente con el fin de poder contrastar los efectos sobre la variable dependiente. Estos efectos se estiman comparando los resultados obtenidos en la variable dependiente.

3) La muestra se elegirá al azar. Si los niveles de la variable independiente se aplican a grupos de sujetos distintos, antes de comenzar el experimento los sujetos se asignan al azar a los grupos donde se aplicarán los niveles de la variable independiente. A cada grupo se le aplica al azar un nivel de la variable independiente. Así se garantiza que por efecto del azar los grupos sean homogéneos o equivalentes, para que después de aplicar los niveles de la variable independiente dichos grupos sean comparables entre sí. De esta manera, las posibles diferencias encontradas al medir la variable dependiente en cada grupo son atribuibles a la aplicación de los niveles de la variable independiente.

5.1 TIPOS DE VARIANZA

La meta ideal del experimentador es que los cambios o variaciones en los valores de la variable dependiente (VD) sean atribuibles a las manipulaciones de la variable independiente (VI) y no a otras variables.

Como ya se ha expuesto, estas variables extrañas podrían influir con intensidad



desigual en los grupos control y experimental(es) y sus efectos se unirían al de la variable independiente, constituyendo fuentes de variación. En consecuencia, el investigador no podría distinguir qué cambios en la variable dependiente son exclusivamente debidos a la variable independiente, y el experimento quedaría invalidado, a menos que dichas variables se controlen igualando o eliminando su influencia. El objetivo del investigador es distinguir los cambios de la variable dependiente que se deben a la manipulación de la variable independiente y los que se deben a otras variables (extrañas). Tanto la variable independiente como las otras variables son fuentes de variación o varianza. Analizaremos las más importantes agrupadas en tres grandes bloques o fuentes de variación:

a) *El experimentador.* Las actitudes y expectativas del investigador sobre los efectos de los niveles de la variable independiente pueden influir en los resultados del experimento. También los rasgos físicos (edad, sexo, aspecto, etc.), la forma de comportarse y las características sociales (status, nivel de estudios, etc.) del experimentador pueden influir.

b) *El contexto experimental.* La posible contaminación de los datos en función del contexto puede provenir de dos fuentes: de cómo captan los sujetos las condiciones de la situación experimental y del diseño (nivel de las variables ambientales, condiciones de aplicación, etc). Veamos brevemente estos aspectos:

- *Demanda de la situación experimental.* La interpretación que los sujetos de los grupos experimentales realizan acerca del experimento puede condicionar su actuación. A través de esta interpretación el sujeto intuye lo que se espera de él (demanda de la situación experimental), y lo puede descubrir durante la realización del experimento, a través de la captación de indicios, y antes del experimento por medio de los rumores sobre la investigación, el lugar donde se realiza y quiénes lo llevan a cabo.

- *Relacionadas con el diseño.* Pueden provenir del medio ambiente (ruido, iluminación, temperatura, etc.) o de las condiciones de aplicación (instrucciones, sistemas de medición, presentación de materiales, estímulos y recogida de datos).

c) *El sujeto experimental.* Cuando los experimentos se llevan a cabo con sujetos universitarios, obligados o voluntarios, hay que cuestionar su representatividad con respecto al resto de la población. Antes de aplicarse los niveles de la variable independiente los grupos pueden diferir en variables como edad, sexo, motivación,

inteligencia o personalidad. Los sujetos de los grupos experimentales pueden verse más motivados por la novedad que los sujetos que pertenecen al grupo control.

Debido a que la acción de estos tres grupos de variables o fuentes de variación es simultánea, cuando se ha elegido una variable independiente, el investigador ha de controlar o igualar la influencia de las restantes fuentes de variación que puedan afectar a la variable dependiente. Procurará que influyan por igual en el grupo control y en el experimental. Si todas las variables extrañas influyen por igual en ambos grupos, las diferencias (varianza) encontradas entre el grupo control y el experimental serán debidas a la variación introducida deliberadamente (niveles de VI).

Supongamos que queremos comprobar la eficacia de un programa basado en la utilización de medios audiovisuales (MA V) para la enseñanza del inglés en primaria. Para ello, seguimos las fases de la aleatorización, y después de comprobar la equivalencia inicial de los grupos (pretest), aplicamos el programa con MA V a un grupo experimental, y a un grupo control le aplicamos el programa sin MA V. Posteriormente, con una prueba paralela al pretest, medimos el rendimiento final en lengua inglesa (postest) y obtenemos las medias especificadas en la tabla 5.1.

TABLA 5.1 Ejemplo de diseño experimental para estudiar la eficacia de los MA V

Grupo	*n*	*Asignación*	*Pretest*	*Programa*	*Postest*
Experimental (GE)	5	Azar	2	Con MAV	9
Control (GC)	5	Azar	2	Sin MAV	5

Con las medias obtenidas podemos efectuar distintas comparaciones. Imaginemos que después de aplicar las pruebas estadísticas adecuadas obtenemos los resultados que aparecen en la tabla 5.2.

TABLA 5.2 Algunas comparaciones que pueden efectuarse en un diseño experimental

Objetivo	*Prueba estadística*	*Medias*	*Decisión*
Equivalencia inicial	t datos independ.	2 - 2	Ha
Equivalencia final (eficacia MA V)	t datos independ.	9 - 5	Hj
Cambio en GE	t datos apareados	9 - 2	H,
Cambio en GC	t datos apareados	2 - 5	H,

Cuando el investigador pretende contrastar una hipótesis a través de un experimento puede proceder del siguiente modo:

1) Mantiene constantes o iguala las variaciones (varianza) que puedan producir todas las posibles influencias que procedan de las variables extrañas. Esto es, controlar la influencia de las variables extrañas.
2) Introduce una variación deliberada controlando (manipulando) la variable independiente.
3) Analiza la intensidad de la variación que se produce en VD como consecuencia de la variación introducida al aplicar distintos niveles de la variable independiente.

De ahí que las conclusiones de un experimento se basen en las variaciones, cambios o varianzas observadas en la variable dependiente a raíz de las variaciones o cambios introducidos deliberadamente en la variable independiente, procurando igualar o sistematizar la varianza debida a variables extrañas. Para entender mejor la importancia de la varianza la analizaremos más detenidamente en el contexto del experimento planteado sobre los MA V.

En la tabla 5.3 se han esquematizado, de forma ficticia y aproximada, algunas fuentes de variación que podrían intervenir en el experimento de los MA V. Las puntuaciones en inglés obtenidas por los sujetos en el pretest presentan ciertas diferencias o variaciones que pueden atribuirse a tres fuentes fundamentales: diferencias individuales (motivación, inteligencia...), influencias ambientales (profesores, familia...) e imperfecciones de los instrumentos de medida. A pesar de esta varianza observada en las puntuaciones de los sujetos los grupos que se han formado tienen la misma media en inglés. Es decir, puede existir cierto grado de variación (debería constatarse estadísticamente) entre los sujetos que forman cada grupo (varianza intragrupo) sin que aparezcan necesariamente diferencias entre los grupos (varianza intergrupos).

Al iniciar el experimento, el investigador introduce una variación: utiliza los MA V en el grupo experimental, pero no en el grupo control. Ahora bien, como puede apreciarse en la tabla 5.3, otras variables como el profesor y la memoria de los sujetos también pueden originar variaciones.

El posible efecto de todas las fuentes de variación mencionadas queda recogido en la varianza total que presentan las puntuaciones en el postest de todos los sujetos. Como puede apreciarse en la tabla 5.3, la variación total de las puntuaciones de inglés se debe a unas variaciones constantes o sistemáticas que constituyen la denominada *varianza sistemática* y a otras irregulares o aleatorias que constituyen la *varianza residual o del error.*

A) *Varianza sistemática.* Es una variación originada por variables que desvían los datos en una dirección más que en otra, aumentando o disminuyendo, pero siempre en la misma magnitud y dirección. Esta desviación, en el experimento perfectamente controlado, es debida exclusivamente a la variable independiente (MA V). Cuando el grado de control no es el ideal, la desviación ha podido ser provocada también por variables extrañas que no han sido convenientemente controladas. Así, en el ejemplo propuesto el entusiasmo del profesor ha influido más en el grupo experimental (+ 2) que en el de control (+ 1).

En cambio, el hecho de que a los sujetos de ambos grupos les hayan proyectado las mismas películas en inglés durante las clases, tiene una incidencia idéntica tanto en el grupo control como en el experimental, es decir, esta influencia extraña ha sido controlada y deja de ser una fuente de variación. Al influir por igual (+ 1) no añade ninguna diferencia o varianza, no afecta a la variación existente entre los datos que resultan de medir VD. Por tanto, si queremos ver la influencia de los MA V debemos controlar las otras fuentes de variación.

TABLA 5.3 Posibles fuentes de variación que pueden intervenir en el experimento de los MA V (Los datos son ficticios; sólo se exponen con fines didácticos para facilitar la comprensión de los conceptos; por tanto, se omite la significación estadística)

	(VD) Pretest	*(VI) MAV*	*Profesor*	*Memoria*	*Error medida*	*Películas*	*(VD) Postest*
GC	2	(+0)	(+1)	(+4)	(-1)	(+1)	7
	1	(+0)	(+1)	(-2)	(+2)	(+1)	3
	3	(+0)	(+1)	(+2)	(-2)	(+1)	5
	2	(+0)	(+1)	(-1)	(+3)	(+1)	6
	2	(+0)	(+1)	(+2)	(-2)	(+1)	4
	$\bar{X} = 2$						$\bar{X} = 5$
GE	1	(+3)	(+2)	(-1)	(+1)	(+1)	7
	4	(+3)	(+2)	(-2)	(-1)	(+1)	11
	1	(+3)	(+2)	(-1)	(+2)	(+1)	10
	2	(+3)	(+2)	(+1)	(-1)	(+1)	8
	2	(+3)	(+2)	(-1)	(+2)	(+1)	9
	$\bar{X} = 2$	Varianza primaria	Varianza secundaria	Varianza intragrupo o del error			$\bar{X} = 9$

↑ desviación sistemática de los datos ↓

Varianza sistemática (Varianza primaria + Varianza secundaria) — Varianza residual (Varianza intragrupo o del error)

El propósito principal del investigador es separar aquellas fuentes de variación en las que se halla interesado (VI) de las que no le interesan (variables extrañas). De ahí la necesidad de considerar dos fuentes de variación sistemática: la varianza primaria y la varianza secundaria.

- *Varianza primaria.* Desviación sistemática de los datos debida a la manipulación de VI. Cuanto más efectiva sea VI, más diferencia o desviación aparecerá entre la medición de VD efectuada en el grupo control y en el experimental. De ahí que se denomine *varianza intergrupos.*
- *Varianza secundaria.* Desviación sistemática de los datos debida a otras variables que no son VI y no han sido controladas. Por ello influyen de forma desigual, provocando varianza no pretendida.

B) *Varianza del error o aleatoria.* Constituida por el conjunto de fluctuaciones que presentan las medidas de la variable dependiente que tienden a actuar al azar. Posibles fuentes son:

- *Diferencias individuales.* Constituidas por características inherentes a los sujetos; por ejemplo, en el experimento de los MA V podría ser la memoria de cada sujeto. Ante un mismo tratamiento experimental no todos los sujetos reaccionan igual. Dado que las diferencias individuales tienden a desviar los datos en sentidos dispares, esta varianza se considera debida al error o aleatoria por ser impredecible.
- *Errores de medida.* Puede deberse a instrumentos poco precisos.
- *Procedimiento experimental.* La forma de aplicarlo puede ser diferente para los distintos sujetos.
-*Instrucciones.* Pueden darse de forma diferente por distintos experu:nentadores.

5.2 PRINCIPIO FUNDAMENTAL DE LA VARIANZA

En el experimento planteado sobre los MA V comprobaremos el siguiente principio fundamental de la varianza, es decir,

Varianza total = Varianza intergrupos + Varianza intragrupos

Como se recordará, los resultados medios de inglés, obtenidos en el postest, eran los siguientes:

Sin MAV		*Con MAV*	
G.C.	Inglés	G	Inglés
a	7	f	7
b	3	9	**11**
e	5	h	10
d	6	i	8
e	4	j	9
	$\bar{X}=5$		$\bar{X}=9$

Sucesivamente procederemos a realizar el cómputo de la varianza total y de las varianzas intergrupo e intragrupo, con el [m de apreciar que aquélla se puede desglosar en la suma de estas últimas.

Cálculo de la varianza total. Variación de cada puntuación con respecto a la media total. Es decir:

$$S^2 = \frac{\sum (X-\bar{X})^2}{n} = \frac{60}{10} = \boxed{6}$$

$$\bar{X} = \frac{\sum X}{n} = \frac{70}{10} = 7$$

	X	$(X-\bar{X})$	$(X-\bar{X})^2$
	7	0	0
Grupo control	3	−4	16
	5	−2	4
	6	−1	1
	4	−3	9
	7	0	0
Grupo experimental	11	4	16
	10	3	9
	8	1	1
	9	2	4
Σ	70		60

Cálculo de la varianza intergrupos (primaria). La varianza total se ha calculado a partir de las diez puntuaciones en lengua inglesa; en cambio, la varianza intergrupo se refiere a la variación entre las medias de los grupos. En consecuencia, se utiliza la misma fórmula, pero en el cálculo sólo intervienen las medias 5 y 9 representando a cada grupo:

$$S^2 = \frac{\sum (X-\bar{X})^2}{n} = \frac{8}{2} = \boxed{4}$$

$$\bar{X} = \frac{\sum X}{n} = \frac{14}{2} = 7$$

	X	$(X-\bar{X})$	$(X-\bar{X})^2$
Grupo control	5	−2	4
Grupo experi.	9	+2	4
Σ	14		8

Cálculo de la varianza intragrupo (del error). Promedio de las variaciones calculadas dentro de cada grupo con respecto a las medias respectivas. El cómputo sería el siguiente:

$$S_c^2 = \frac{\sum (X-\bar{X})^2}{n} = \frac{10}{5} = \boxed{2}$$

$$\bar{X} = \frac{\sum x}{n} = \frac{25}{5} = 5$$

	X	$(X-\bar{X})$	$(X-\bar{X})^2$
	7	−2	4
Grupo control	3	−2	4
	5	0	0
	6	+1	1
	4	−1	1
Σ	25		10

De esta manera queda comprobado que se cumple el principio fundamental de la varianza:

$$S_e^2 = \frac{\Sigma (X-\bar{X})^2}{n} = \frac{10}{5} = 2$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{45}{5} = 9$$

	X	$(X-\bar{X})$	$(X-\bar{X})^2$
	7	−2	4
Grupo	11	+2	4
experi-	10	+1	1
mental	8	−1	1
	9	0	0
Σ	45		10

$$S^2 = \frac{S_c^2 + S_e^2}{n} = \frac{2+2}{2} = \boxed{2}$$

De esta manera queda comprobado que se cumple el principio fundamental de la varianza:

6	=	4	+	2
Varianza total	=	Varianza intergrupos	+	Varianza intragrupos
Varianza total	=	Varianza sistemática	+	Varianza del error
Varianza total	=	Varianza primaria (VI) + Varianza secundaria (Extrañas)	+	Varianza residual (Dif. individuales errores)

Las pruebas de significación, como el análisis de la varianza, se basan en el cociente:

$$F = \frac{\text{Varianza intergrupos}}{\text{Varianza intragrupos}} = \frac{4}{2} = 2$$

En la medida que aumente la varianza intergrupos (refleja el efecto de VI) y disminuya la varianza intragrupos (diferencias individuales y errores) el experimento será más eficaz y sensible, pues el cociente aumentará. Si nuestro experimento ficticio estuviera perfectamente controlado se verificaría que:

varianza secundaria = 0

pues las variables extrañas influirán por igual en cada grupo o se habrá eliminado su influencia. Por tanto, en el experimento ideal,

varianza intergrupos = varianza primaria = 4 (Sistemática) (VI)

Por tanto, en el buen experimento:

varianza intergrupo > varianza intragrupo

Para ello es necesario un control que consistirá en:

1) Maximizar la varianza primaria (producida por VI).
2) Neutralizar o eliminar la varianza secundaria (producida por las variables extrañas).
3) Minimizar la varianza del error (producida por las diferencias individuales y los errores de medida).

5.3 CONTROL DE LA V ARIANZA

El experimento es un tipo especial de investigación para examinar relaciones de causalidad. Para ello se introducen cambios deliberados en la variable independiente y se registran posibles modificaciones en la variable dependiente.

Con el fin de garantizar la validez del experimento, se utilizan técnicas de control de la varianza que pueden consistir en manipulación física, selectiva o estadística:

Manipulación física. Son técnicas de control directo que pueden ejercerse al aplicar la variable independiente o al igualar los efectos de algunas variables extrañas. Por ejemplo, espejos unidireccionales, suministrar drogas, premios, presentar a los sujetos estímulos verbales, luminosidad, ruidos o temperatura.

Manipulación selectiva. Algunos ejemplos pueden ser seleccionar y asignar aleatoriamente los sujetos, mantener constantes ciertas condiciones, aparear los sujetos, el bloqueo, el balanceo y el contrabalanceo.

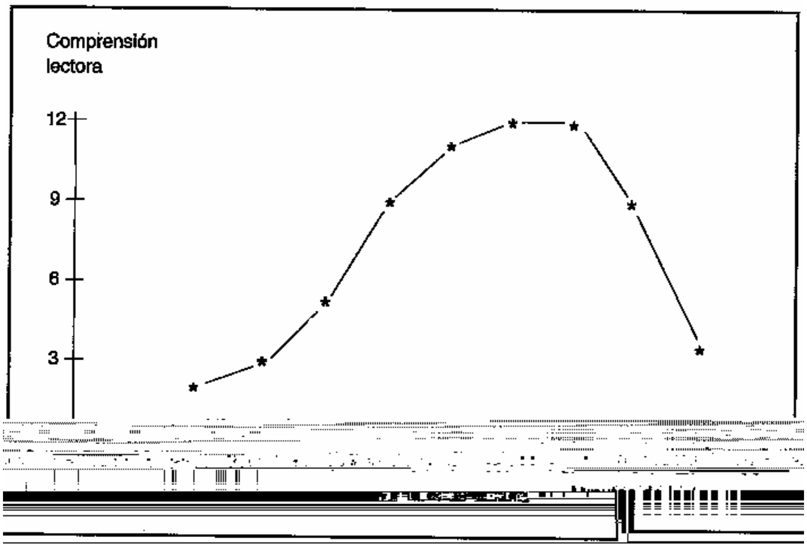
Manipulación estadística. Dada la naturaleza del fenómeno educativo es frecuente tener que recurrir a técnicas indirectas de control como son las técnicas estadísticas. Estas técnicas de control, como la correlación parcial y el análisis de covarianza, se utilizan cuando las variables no pueden ser sometidas a manipulación física o selectiva.

Como ya hemos apuntado más arriba, el investigador ha de recurrir a estas técnicas de control con el fin de maximizar la varianza primaria, eliminar la varianza secundaria

y minimizar la varianza del error. A continuación veremos más detenidamente distintas técnicas de control:

MAXIMIZACIÓN DE LA V ARIANZA PRIMARIA

Los niveles de la variable independiente han de adoptar en la realidad los valores decididos por el experimentador. Además, dichos valores han de ser extremos y óptimos. Supongamos que se realizan distintos experimentos en tomo a la eficacia de técnicas de estudio en niños de nueve años. Si queremos ver la influencia del tiempo de lectura en el rendimiento, habrá que asignar los valores adecuados. Así, valores como O, 1 Y 2 minutos de lectura pueden llevamos a concluir que el tiempo de lectura no influye, ya que no son valores extremos ni óptimos. Lo mismo podría ocurrir con valores más extremos, pero no óptimos, como O, 50 y 90 minutos, ya que al visualizar la relación entre dicha variable y una característica decisiva para el rendimiento (comprensión) se obtiene la figura 5.1.



FIG, 5.1. Distribución de la comprensión lectora según los minutos de lectura

En cambio, valores como O, 25 y 50 podrían ejercer un efecto más diferenciado, caso de existir una relación real. En consecuencia, podrían ser valores extremos y óptimos.

Existen tres formas de asignar los niveles de la variable independiente a los sujetos que permiten distintas modalidades de control: intergrupo, intragrupo y mixta.

Asignación intergrupo

Se utiliza esta modalidad cuando cada nivel de la variable independiente, también denominado *condición experimental* o *tratamiento,* se aplica a un grupo diferente de sujetos, es decir, cada sujeto es sometido a una sola condición experimental. Así, por ejemplo, para ver cómo influye el tiempo de estudio en la comprensión lectora de un texto T podemos seleccionar al azar 30 alumnos de B UP y asignarlos, también al azar, a un grupo control (GC) y dos grupos experimentales (GEl y GEl). Los alumnos asignados al GC no dedican tiempo a estudiar el texto, mientras que en cada GE se dedica 1 y 2 horas. Después de realizar el proceso descrito se comparan las medias de comprensión lectora referidas a dicho texto (tabla 5.4).

TABLA 5.4 Ejemplo de asignación de sujetos según la modalidad intergrupo

Grupo control (O horas)	*Grupo experimental* 1 (1 hora)	*Grupo experimental* 2 (2 horas)
1.Juan	11. Rosa	21. Lourdes
2. Ana	12. José	22. Carlos
3. Antonio	13. Raquel	23. Miguel
………….....	...	...
9. Ramón	19. Mireia	29. Sergio
10. María	20. Víctor	30. Mónica
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$

En este experimento, comparar las medias tiene sentido si los tres grupos que intervienen y las tres situaciones ya eran homogéneas o equivalentes, antes de empezar el experimento y durante el desarrollo del mismo, especialmente en lo que se refiere a diferencias que pueden ser ocasionadas por variables extrañas que pueden influir en la variable dependiente. Para conseguir que estas variables afecten por igual a los tres grupos de sujetos y hacerlos comparables hay que controlar o igualar en cada grupo influencias como las siguientes:

- El experimentador: entusiasmo, amabilidad, etc.
- El contexto experimental: ruido, local, hora, instrucciones, etc.
- Los sujetos experimentales que presentan diferencias individuales en motivación, inteligencia, edad, sexo, capacidad de estudio, etc.

Se consigue un control ideal cuando lo único que diferencia a cada grupo es el tratamiento proporcionado a través de los niveles de la variable independiente (O horas, 1 hora y 2 horas). En consecuencia, si antes de empezar el experimento los tres grupos presentaban medias estadísticamente iguales en la comprensión lectora del texto T,

puede concluirse, con más seguridad, que si aparece alguna diferencia entre las medias después del experimento, dicha diferencia podrá ser atribuida a que los grupos han sido sometidos a distintos niveles de la variable independiente.

Asignación intragrupo

Según esta modalidad todos los niveles de la variable independiente se aplican sucesivamente al mismo grupo de sujetos, es decir, cada sujeto es sometido a todas las condiciones experimentales. En el ejemplo anterior, podemos someter a los 30 sujetos a la prueba de comprensión lectora sobre el texto T para obtener la media cuando el nivel de la variable independiente vale O horas de estudio. En dos días sucesivos el grupo total estudia durante 1 hora y 2 horas respectivamente, midiéndose la comprensión después de cada período de estudio (tabla 5.5).

TABLA 5.5 Ejemplo de asignación de sujetos según la modalidad intragrupo

Grupo total (0 horas)	*Grupo total* (1 hora)	*Grupo total* (2 horas)
1. Juan	1. Juan	1. Juan
2. Ana	2. Ana	2. Ana
3. Antonio	3. Antonio	3. Antonio
...	...	...
29. Sergio	29. Sergio	29. Sergio
30. Mónica	30. Mónica	30. Mónica

–

Aunque existen técnicas de control que tratan de igualar la influencia de estas variables extrañas, para el experimento planteado es más viable una aplicación intergrupo; sin embargo, el mismo ejemplo ha sido incluido aquí con fines meramente expositivos.

Asignación mixta

En esta modalidad a varios sujetos se les aplica solamente un nivel de la variable independiente (intergrupos) y a otros sujetos se les aplican varios niveles de la variable independiente (intragrupo). Así, en el ejemplo mencionado, después de aplicar la prueba de comprensión al grupo total, el mismo grupo de 30 sujetos puede estudiar durante 1 hora, mientras que sólo 10 sujetos estudian durante 2 horas (tabla 5.6).

TABLA 5.6. Ejemplo de asignación de sujetos según la modalidad mixta

Grupo total (O horas)	*Grupo total* (1 hora)	*Grupo total* (2 horas)
1. Juan	1. Juan	1. Rosa
2. Ana	2. Ana	2. José
3. Antonio	3. Antonio	3. Raquel
...	...	...
...		
29. Sergio	29. Sergio	29. Mireia
30. Mónica	30. Mónica	30. Víctor
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$

Esta manera de asignar los tratamientos combina las ventajas de las condiciones intergrupo e intragrupo.

ELIMINACIÓN DE LA V ARIANZA SECUNDARIA

El experimentador ha de garantizar que los cambios observados en la variable dependiente (varianza) sean únicamente debidos a los distintos niveles de la variable independiente. Para ello las variables extrañas no han de contaminar los resultados finales del experimento, es decir, las variables extrañas han de incidir con la misma intensidad en los grupos experimentales y en el grupo control. De esta forma, los grupos experimentales y el grupo control sólo se diferenciarían en que reciben distintos niveles de la variable independiente. En consecuencia, las diferencias observadas en la variable dependiente serán atribuibles a los niveles de la variable independiente siempre que las variables extrañas no influyan de forma diferenciada.

Si el experimento no se controla adecuadamente, existe la posibilidad de que los

cambios observados en la variable dependiente (varianza total) puedan ser provocados, en parte, por una gran multiplicidad de variables extrañas que junto a la variable independiente pueden producir modificaciones en la variable dependiente. Las variables extrañas que pueden afectar a los resultados finales han sido englobadas en categorías como las siguientes:

- Sujetos experimentales: edad, sexo, capacidad cognitiva, personalidad, destreza, etc.
- Contexto experimental: instrucciones, aparatos, ruidos, etc.
- Experimentador: expectativas, actitudes, *rol,* etc.

De ahí que el investigador deba cerciorarse de que las variaciones que aparecen en la variable dependiente puedan ser atribuidas exclusivamente a las variaciones introducidas deliberadamente a través de la variable independiente. Para ello puede utilizar distintas técnicas de control.

Las principales formas y técnicas de *control de las variables extrañas* se basan fundamentalmente en la *aleatorización* y la selección. La *aleatorización* constituye el procedimiento más adecuado para controlar las variables extrañas conocidas y desconocidas que proceden de los sujetos. Gracias a la *selección* un mismo método didáctico puede utilizarse para explicar a los alumnos diferentes conceptos que pueden seleccionarse según su dificultad o adecuación a la edad. Es posible formar grupos según el sexo, edad, clase social, motivación, inteligencia o personalidad cuando sospechamos o sabemos que estas variables pueden tener una influencia significativa. Mediante la selección se pueden eliminar variables, como pueden ser el ruido, o bien elegir un valor único o una única categoría: niños o niñas, período evolutivo, procedencia, etc., teniendo en cuenta que limitaremos el poder generalizador de las conclusiones. Algunas variables contextuales, como la temperatura y la hora de aplicación de los instrumentos de medida, también podemos mantenerlas constantes.

A continuación se detalla la técnica de la aleatorización y diversos procedimientos, algunos de ellos basados en la selección:

Aleatorización

Consiste en asignar los sujetos al azar a cada grupo del experimento, lo que permite igualar todas las diferencias individuales (carácter, inteligencia, motivación, comprensión lectora, etc.). La asignación aleatoria asegura que dichas variables tengan la misma oportunidad de influir en la variable dependiente, llegando a compensarse sus efectos y resultando grupos equivalentes antes de aplicarse la variable independiente.

La aplicación estricta de esta técnica requiere las siguientes fases (fig. 5.2):

1) Seleccionar al azar una muestra de la población.

2) Asignar los sujetos al azar a tantos grupos como niveles tenga la variable independiente. Por ejemplo, grupo 1 (G_1) y grupo 2 (G_2).

3) Formados los grupos del experimento se asigna al azar cada nivel de la variable independiente a un grupo. El grupo que reciba el nivel mínimo de la variable independiente o ausencia de tratamiento experimental (T_o), se denominará grupo control

(GC), Y grupo experimental (GE) el que reciba el tratamiento (TI). De ahí que estas dos últimas fases sean más propias de la modalidad de asignación intersujetos.

Cuando el número de sujetos de la muestra es suficientemente grande, la probabilidad de que las diferencias individuales se compensen unas a otras, por efecto del azar, es también mayor, existiendo más garantía de que los grupos sean equivalentes antes de iniciarse el experimento. Como es evidente, aquellas variables extrañas que puedan influir o generarse después de haberse formado los grupos (cansancio, ansiedad, experimentador, ambiente, contexto experimental, etc.) no pueden controlarse mediante la técnica de la aleatorización.

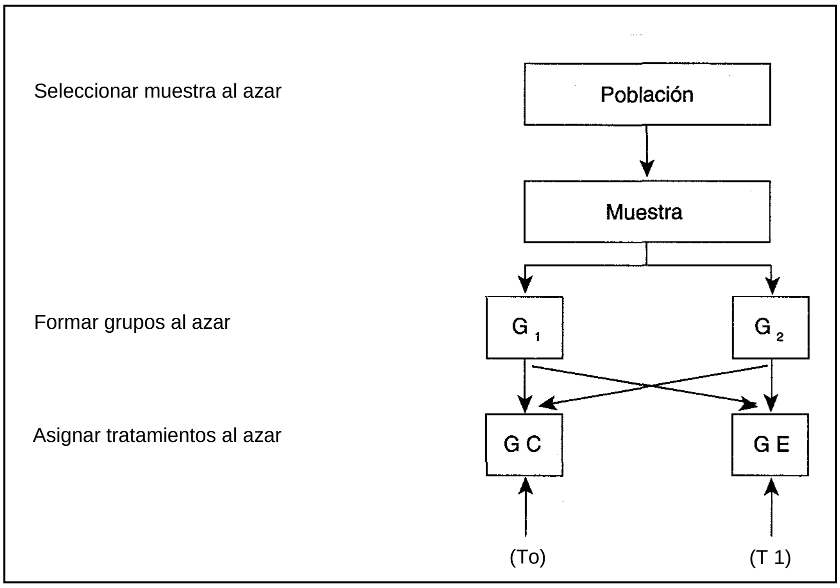


FIG. 5. 2 Fases de la aleatorización como técnica de control (adaptado de Rosel, 1986,244)

Constancia

Las variables extrañas se mantienen a un nivel constante. Por ejemplo, si queremos ver la eficacia de distintos métodos didácticos se procurará que para todos los grupos del experimento los temas que hay que explicar sean los mismos. También convendrá que otras características, como el nivel de iluminación del local, la claridad al especificar las instrucciones, etc., intervengan por igual en cada grupo del experimento. La constancia (y también la eliminación) presenta el inconveniente de que restringe la generalización de las conclusiones del experimento. Así, cuando la edad se mantiene constante a los 13 años o cuando sólo se seleccionan chicos, los resultados del experimento no serán extensibles a los alumnos de 10 años ni a las chicas, respectivamente.

Eliminación

Es un caso especial de constancia que se lleva a cabo cuando algunas condiciones ambientales se restringen al máximo, manteniéndolas constantes pero a nivel de ausencia, es decir, a nivel cero. Por ejemplo, pueden eliminarse elementos distorsionantes como el ruido, la visión de objetos que distraen, etc., lo que hace que ignoremos qué ocurriría si el experimento se llevara a cabo en condiciones normales.

Apareo

Cuando sospechamos que una variable extraña (variable de apareo) está correlacionada y/o puede influir decisivamente en la variable dependiente, podemos formar grupos equivalentes antes de aplicar la variable independiente, recurriendo al apareo de los sujetos. Para ello asignamos a cada grupo del experimento un sujeto que posea la misma magnitud en la variable de apareo o de bloqueo.

Supongamos que queremos comprobar la eficacia de un programa de informática para resolver ecuaciones. Podemos utilizar la inteligencia como variable de apareo, después de comprobar que se halla fuertemente correlacionada con la habilidad para resolver ecuaciones. En la práctica, aunque se utilice la aleatorización, cuando la muestra no es muy grande, puede darse el caso de que el grupo experimental esté constituido por sujetos más inteligentes. Si así ocurriera, el posible aumento en la habilidad para resolver las ecuaciones podría deberse no sólo al programa de informática, sino también a la inteligencia. Así, para aplicar la técnica del apareo en el ejemplo propuesto podríamos aplicar el test de inteligencia «D-48», lo que nos permitiría formar un grupo control (GC) y otro experimental (GE), agrupando por parejas los sujetos más afines en dicha variable y asignándolos al azar a cada grupo, lo que origina distintos pares o bloques (tablas 5.7 y 5.8).

Como los sujetos dentro de cada bloque o par tienen la misma inteligencia, resultará que los grupos control y experimental serán equivalentes antes de aplicar el programa de informática. Como es evidente, la propia variable dependiente también podría utilizarse como variable de apareo.

TABLA 5. 7 Puntuaciones de inteligencia «D-48»

1. Eva	30	7. Ignacio	35
2. Pablo	33	8. Rosa	28
3. Carlos	34	9. Javier	33
4. Gemma	35	10. Ana	30
5. Manuel	28	11. Marco	27
6. Carmen	27	12. Sonia	34

TABLA 5.8 Pares de sujetos y asignación a cada grupo

Pares	GC	*GE*
1.	Ignacio	Gemma
2.	Carlos	Sonia
3.	Pablo	Javier
4.	Eva	Ana
5.	Rosa	Manuel
6.	Carmen	Marco
	X_0	X_j

En la representación simbólica de un diseño basado en la técnica del apareo los distintos pares pueden representarse por:

Ap = asignación según apareo

En general, esta técnica de control presenta inconvenientes como los siguientes:

- Cuando la variable de apareamiento presenta gran variabilidad hay que tener más flexibilidad en el grado de semejanza que han de tener las puntuaciones para incluidas en el mismo par.
- Con frecuencia, hay que prescindir de algunos sujetos si no tienen un «par» adecuado que presente un nivel similar en la variable de apareamiento.
- De la misma manera, cuando ya se han formado los pares y falla un sujeto por enfermedad o absentismo, hay que eliminar el par si no existe otro sujeto similar que pueda sustituir al ausente.
- Si se quiere aparear a los sujetos a partir de más de una variable de apareo, las posibilidades de formar los pares son menores y la muestra quedaría muy reducida.

Bloqueo

Cuando se forman más de dos grupos de tratamiento, en lugar de agrupar por pares interesa formar bloques a partir de una o más variables de bloqueo. Esta técnica presenta inconvenientes similares a los que se han expuesto al considerar la técnica del apareo y hay que admitir un mayor margen de variabilidad en los sujetos que se agruparán en un mismo bloque; de lo contrario habría que prescindir de muchos sujetos. Sin embargo, como dentro de cada bloque los sujetos se asignan también al azar a cada grupo de tratamiento, es fácil conseguir grupos equivalentes con respecto a la variable de bloqueo.

En el ejemplo planteado en el apartado anterior, si queremos comprobar la eficacia relativa de tres programas de informática para resolver las ecuaciones, podemos formar un grupo control y tres grupos experimentales, utilizando las puntuaciones de inteligencia para agrupar tres bloques de inteligencia alta, media y baja (tabla 5.9).

Balanceo

Esta técnica distribuye equitativamente la presencia de una variable extraña en los grupos que deseamos formar. Si en el experimento mencionado anteriormente sobre la eficacia de los programas de informática se sospecha que las chicas pueden tener una mayor habilidad para la resolución de ecuaciones, no sabríamos si la posible mejora observada en los grupos experimentales es debida a los programas, al hecho de que en dichos grupos pudiera predominar el sexo femenino o a ambas circunstancias.

TABLA 5.9 Ejemplo ilustrativo de la técnica de bloqueo

Bloques autorizados	*GC*	*GE₁*	*GE₂*	*GE₃*
1. Inteligencia alta	Ignacio	Carlos	Gemma	Sonia
2. Inteligencia media	Eva	Javier	Ana	Pablo
3. Inteligencia baja	Carmen	Marco	Rosa	Manuel
	$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$

Supongamos que al comparar las medias se obtiene que bajo el programa 2 (GE2) la habilidad para resolver las ecuaciones es superior a la que presentan los sujetos que han utilizado el programa 1 (GE1). Esta diferencia entre las medias obtenidas podría deberse a que el programa 2 es más eficaz que el programa 1 y/o al hecho de que el grupo GE2 está constituido por chicas y el grupo GE1 por chicos. En consecuencia, los resultados del experimento quedarían contaminados por la variable sexo. Hay que tener en cuenta que la muestra debería ser más grande y entonces sería menos probable que predominara uno de los sexos en algún grupo después de aleatorizar la asignación dentro de cada bloque. Una posible solución sería formar los grupos incluyendo la misma cantidad de chicos y chicas en cada grupo. Para recoger la mayor información posible con respecto a los efectos de la variable sexo, y suponiendo que no interese controlar la inteligencia, el experimento podría diseñarse según la técnica del balanceo para controlar la variable sexo. Para una mayor sencillez veamos cómo podría comprobarse la eficacia de un solo programa (tabla 5.10).

TABLA 5. 10 Ejemplo con la utilización de la técnica del balanceo

		GC	GE	
Sexo	M	Ignacio Carlos Pablo	Javier Marco Manuel	$\bar{X}_M$
	F	Eva Carmen Gemma	Ana Rosa Sonia	$\bar{X}_F$
		$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_T$

Como puede observarse, el experimento tiene cuatro subgrupos de sujetos: dos subgrupos forman el grupo control (tres chicos y tres chicas) y el programa es aplicado a otros dos subgrupos (tres chicos y tres chisas) y dos subgrupos. Si se incluyen más sujetos en cada subgrupo, la comparación $\bar{X}_0$ y $\bar{X}_1$ nos informa sobre el efecto del programa, ya que en el grupo control no ha sido aplicado. Además, podemos obtener una información adicional sobre el sexo al comparar las medias $\bar{X}_M$ y $\bar{X}_F$. En realidad, hemos sistematizado o equilibrado una variable extraña (sexo), considerándola como otra variable independiente de la investigación.

Contrabalanceo o equiponderación

Con esta técnica, los grupos se dividen en pequeños sub grupos con el fin de equilibrar los efectos de las variables extrañas. Puede utilizarse en la modalidad de asignación intragrupo, es decir, cuando distintos tratamientos aplicados sobre los mismos sujetos pueden producir efectos distintos según la secuencia u orden de aplicación. Supongamos que hemos elaborado varias reglas nemotécnicas para potenciar las facultades y alcance de la memoria de un grupo de niños débiles mentales y pretendemos comprobar la eficacia relativa de dos técnicas (T) y (T) para memorizar frases. Como es obvio, al aplicar cada técnica (VI) y al contabilizar cuántas frases han memorizado (VD), se recurriría a frases distintas pero de dificultad similar. Sin embargo, en cada aplicación habría que tener en cuenta dos variables extrañas que pueden generarse durante el experimento:

1) Cierto nivel de aprendizaje práctico, experiencia o habilidad que puede facilitar la memorización de las frases después de aplicar la primera técnica.
2) La fatiga, que puede favorecer que el sujeto cometa más errores a medida que aumenta el número de tratamientos.

El grado de memorización de las frases dependerá no sólo de la eficacia de la técnica aplicada, sino también de la acción conjunta del aprendizaje y de la fatiga, que originan el denominado *error progresivo.* Aunque éste puede disminuir con adecuados intervalos intertratamientos, ya que el tiempo suele eliminar los efectos de dicho error progresivo, es más viable neutralizarlo por medio del contrabalanceo (tabla 5.11).

TABLA 5. 11 Técnica de contrabalanceo. Modalidad intragrupo

T_1 (1.ª)	T_2 (2.ª)	T_2 (3.ª)	T_1 (4.ª)
Alicia	Alicia	Javier	Javier
Iván	Iván	Isabel	Isabel
Roberto	Roberto	Laura	Laura
...	...	...	...
...	...	...	...
Olga	Olga	Irene	Irene
Ramón	Ramón	Vicente	Vicente
$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$

La equiponderación o contrabalanceo puede adoptar también la modalidad intergrupo (tabla 5. 12).

TABLA 5. 12 Técnica de contrabalanceo. Modalidad intergrupo

T_1 (1.ª)	T_2 (2.ª)	T_2 (3.ª)	T_1 (4.ª)
Alicia	Alicia	Javier	Javier
Iván	Iván	Isabel	Isabel
Roberto	Roberto	Laura	Laura
...	...	...	...
...	...	...	...
Olga	Olga	Irene	Irene
Ramón	Ramón	Vicente	Vicente
$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$

En ambas modalidades se ha incluido entre paréntesis la secuencia de aplicaciones y la media del número de frases memorizadas después de haber practicado con cada técnica. Así, los niveles medios que permitirán estimar la eficacia relativa de cada regla nemotécnica podrían ser:

$$X_{T1} = \frac{\bar{X}_1 + \bar{X}_2}{2} \qquad X_{T2} = \frac{\bar{X}_3 + \bar{X}_4}{2}$$

ANÁLISIS DE COV ARIANZ

Las técnicas de control descritas pueden plantear al educador dificultades como las siguientes:

1) Formar nuevos grupos para el experimento, por ejemplo, a través de la aleatorización, puede ser no recomendable por razones organizativas de los centros. Esto obliga a utilizar grupos naturales, respetando las agrupaciones que constituyen las distintas clases o conglomerados.
2) La utilización de otras técnicas como el apareo y los bloques puede resultar inapropiada, ya que el educador suele disponer de pocos sujetos para formar grupos homogéneos igualados en variables extrañas relevantes.

En consecuencia, al tomar grupos ya constituidos (naturales), las variables extrañas pueden estar presentes con distintos niveles de intensidad, originando diferencias entre los grupos antes de empezar el experimento. Si estas diferencias iniciales no se pudieran controlar (igualar) mediante algún sistema alternativo de control, la investigación no podría realizarse. Por ejemplo, podría darse la situación ficticia de la tabla

5.13 al aplicar dos métodos de enseñanza (M_1 y M_2) a dos grupos de tercero de primaria y analizar sus efectos sobre el rendimiento tomando un grupo control (GC).

TABLA 5.13 Datos de una investigación en la que no se ha controlado la inteligencia

Grupos naturales	*Número sujetos*	*Inteligencia media (0-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
3.ºA	26	30	2	GC	4
3.º B	29	28	1	M_1	7
3.º C	27	32	3	M_2	8

Supongamos que la inteligencia está significativamente correlacionada con el rendimiento, y no podemos controlarla mediante las técnicas mencionadas. El análisis de covarianza permite un control estadístico o control indirecto de esta variable extraña al tener en cuenta las diferencias iniciales entre tercero A, tercero B y tercero C. Gracias al análisis de covarianza el experimento puede realizarse. Esta técnica estadística ajusta las diferencias finales con respecto a las iniciales y compara las medias tomando en consideración la posible existencia de diferencias iniciales que podrían enmascararlas. Es decir, permite corregir o «ajustar» los cambios observados en la variable dependiente, dejando solamente los efectos atribuibles a la variable independiente.

La situación planteada es comparable a lo que ocurriría con tres niños en un concurso para coleccionar cromos. Si el educador quiere saber quién es más hábil coleccionando y sospecha que los niños empiezan el concurso con distinto número de cromos, habrá de comparar las cantidades finales acumuladas tomando en consideración la cantidad inicial que ya tenía cada niño. Es decir, ha de «ajustar» o hacer comparables las cantidades finales. La pregunta implícita sería: ¿cuántos cromos habría ganado cada niño si hubieran empezado sin ninguno o con la misma cantidad?

En el caso que nos ocupa el análisis de covarianza resuelve estadísticamente este interrogante y compara las medias que tendría después cada grupo si inicialmente las hubieran tenido iguales. Por ejemplo, las medias definitivas comparables o ajustadas hubieran podido ser 4, 8 y 7, lo que podría sugerir cierta supremacía del método M 1 (tabla 5.14):

TABLA 5. 14 Datos ficticios de una investigación en la que se ha controlado la inteligencia ajustando las medias a través de un análisis de covarianza

Grupos naturales	*Número sujetos*	*Inteligencia media (0-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
6.º A EGB	26	30	2	GC	4
6.º B EGB	29	28	1	M_1	8
6.º C EGB	27	32	3	M_2	7

En cierto modo, la aportación del análisis de covarianza sería similar a la información que ofrecería el siguiente diseño si en el caso planteado se hubiera podido controlar la inteligencia aleatorizando o seleccionando grupos de sujetos equivalentes (tabla 5.15).

TABLA 5.15 Datos de una investigación en la que se ha controlado la inteligencia directamente aleatorizando o seleccionando grupos equivalentes

Grupos naturales	*Número sujetos*	*Inteligencia media (D-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
3.º A EGB	26	30	2	GC	4
3.º B EGB	29	30	2	M,	8
3.º C EGB	27	30	2	M_2	7

La variable extraña (inteligencia) que se distribuye de forma desigual en los grupos del experimento recibe el nombre de *covariable.* El análisis de covarianza se basa en el supuesto de que el nivel en rendimiento obtenido por cada sujeto depende del valor que tenga en inteligencia y del método que se le aplique. En consecuencia, la media en rendimiento se reajusta en función de los métodos de enseñanza y en función de la covariable inteligencia.

MINIMIZACIÓN DE LA V ARIANZA INTRAGRUPO

Como se recordará, la varianza del error, residual o intragrupo se refería a variaciones debidas al azar. También habíamos contemplado algunas fuentes de esta varianza, como son las diferencias individuales y los errores de medida debidos a la imprecisión de los instrumentos utilizados en el proceso de medición.

Para minimizar la varianza del error conviene utilizar instrumentos fiables al medir la variable dependiente, reducir las diferencias individuales, aplicar bien los niveles de la variable independiente y comunicar las mismas instrucciones a los sujetos que intervienen en el experimento.

5.4 FUENTES DE INVALIDEZ

Hemos visto que en las investigaciones hay que tener en cuenta tres tipos de validez: la interna, que exige el control; la externa (generalización) y la conceptual (coherencia). En el caso de la investigación experimental estos tipos de validez revisten una importancia decisiva. Las variables o elementos que ponen en peligro la validez del experimento se denominan *fuentes de invalidez.* Por su relevancia nos centraremos en las fuentes de invalidez interna y externa.

FUENTES DE INVALIDEZ INTERNA

Cuando un experimento carece de validez interna, el investigador no tiene garantías de que las diferencias observadas (varianza total) en la variable dependiente puedan atribuirse exclusivamente a la aplicación intencional de los distintos niveles de la variable independiente. Existe la posibilidad de que dicha variación sea debida a la influencia de otras variables extrañas o fuentes de invalidez, es decir, hay motivos para sospechar que las variables extrañas no influyen por igual en todos los grupos del experimento, ya que no se aplican técnicas de control adecuadas. En consecuencia, cualquier elemento que afecte al grado de control del diseño de investigación constituye una fuente potencial de invalidez interna.

Las variables extrañas, que de no controlarse pueden constituir fuentes de invalidez interna, las agruparemos bajo las siguientes denominaciones genéricas:

a) *Historia.* Engloba cualquier acontecimiento externo a los sujetos y diferente de la variable independiente que influya a lo largo del experimento (entre el pretest y el postest) en la variable dependiente. Estos elementos distorsionantes pueden producirse dentro o fuera del marco del experimento. El efecto puede agudizarse cuando

1) el tratamiento es largo;
2) se trata de un experimento de campo;
3) los grupos no reciben los tratamientos simultáneamente.

b) *Maduración.* Consiste en un cambio que se produce en los sujetos debido a su desarrollo biológico o psicológico al transcurrir el tiempo. Puede estimarse a través del cambio observado en un grupo control.

c) *Selección diferencial de los sujetos.* Puede ocurrir en investigaciones cuasiexperimentales, cuando no se forman al azar los distintos grupos que intervienen en el diseño. Así, por ejemplo, cuando el investigador ha de utilizar grupos ya constituidos como aulas de un colegio. Dado que en cada grupo las variables extrañas han podido influir previamente de forma diferenciada, no puede garantizarse que los grupos sean homogéneos antes de aplicar la variable independiente. En consecuencia, los cambios observados pueden ser debidos a las diferencias existentes entre los grupos más que a la acción del tratamiento.

d) *Pérdida de sujetos.* Esta fuente de invalidez, también denominada *mortalidad experimental,* puede afectar especialmente si la pérdida se produce en los sujetos situados en los extremos de la distribución.

e) *Interacciones.* Pueden aparecer entre selección y maduración, selección e historia, etc. El grupo experimental puede hallarse más motivado, o ser más ágil al entender las instrucciones, y por ello, independientemente del tratamiento, puede obtener mayor puntuación que el grupo control en el postest.

f) *Reactividad de la medida.* Las medidas o la aplicación de pruebas puede afectar a los sujetos, o pueden responder éstos de forma diferente a como lo harían en condiciones normales, lo que influiría en los resultados finales, independientemente de la acción de la variable independiente.

g) *Instrumentación.* Ciertos instrumentos de medida pueden cambiar con el tiempo. Por ejemplo, algunos sujetos pueden hacer marcas en los protocolos de

algunas pruebas que condicionarán las respuestas de otros sujetos si se utilizan los mismos protocolos, o un observador puede cambiar de criterio durante la recogida de datos o durante el análisis de los mismos. También puede ocurrir que los instrumentos de medida no tengan una fiabilidad adecuada.

h) *Regresión estadística.* Cuando se aplica una prueba en sucesivas ocasiones (pretest-postest) las puntuaciones progresivas tienden a aproximarse a la media del grupo de procedencia. Este efecto es más pronunciado cuando se seleccionan grupos en base a puntuaciones extremas. Por ejemplo, si queremos probar la eficacia de un programa de técnicas de estudio para los mejores alumnos de un curso de primaria podemos aplicar un test de rendimiento con el fin de seleccionar los alumnos que presentan las puntuaciones más altas. Si se les pasa posteriormente un nuevo test de rendimiento, independientemente de la aplicación de las técnicas de estudio, la media que obtendrá el grupo en este segundo test tenderá a desplazarse hacia la media de la población original, es decir, tiende a descender. Si se hubieran seleccionado los alumnos más retrasados, la segunda vez que se pasara el test la media tendería a incrementarse. En ambos casos ocurre un fenómeno de regresión estadística, principalmente debido al azar y a posibles deficiencias de los instrumentos.

La regresión estadística produciría el efecto que aparece en la figura 5.3. Supongamos que al aplicar el primer test el grupo total obtiene una media global de 5, el grupo de alumnos más retrasado sólo registra una media de 3 y los más adelantados de 8. Si realizamos la segunda aplicación de un test de rendimiento, cubriendo el mismo contenido y sin introducir ningún cambio en los sujetos, la media del grupo más retrasado tenderá a aumentar, aproximándose a la media global del grupo de procedencia (5). Por el contrario, la media del grupo más adelantado tenderá a descender aproximándose a la misma media global. Así, si un grupo con puntuaciones altas en alguna variable se utiliza como grupo control, y otro con puntuaciones bajas se constituye en experimental, la regresión no está controlada y puede actuar favoreciendo a los sujetos experimentales.

i) *Difusión del tratamiento.* Los sujetos del grupo experimental pueden comunicar a los demás las características del tratamiento.

FUENTES DE INVALIDEZ EXTERNA

Las fuentes de invalidez externa son condiciones que afectan a la representatividad o generalización de los resultados de una investigación, es decir, a la posibilidad de generalizar los resultados a otros sujetos, a otros grupos y a otras condiciones y situaciones reales.

Las fuentes de invalidez externa más importantes que pueden afectar a la representatividad de los resultados son:

a) *Reactividad experimental.* Cuando los sujetos sometidos a las condiciones experimentales actúan de diferente forma de como lo hacen en condiciones normales. Esta variable puede afectar también a la validez interna.

b) *Interacciones.* Pueden producirse entre los errores de selección y la variable experimental. Cuando se realiza un experimento sobre una muestra sesgada (voluntarios, sujetos de un determinado nivel cultural, social, etc.), los resultados no podrán

generalizarse a otros grupos diferentes no representados en la muestra. De la misma manera, si el experimento se lleva a cabo en un contexto muy específico (laboratorio, colegio, centro de rehabilitación, etc.), no existen suficientes garantías de que los resultados puedan extrapolarse a otros ámbitos. En general, para que los resultados puedan generalizarse a situaciones de la vida real, el experimento ha de realizarse en condiciones similares a contextos reales.

c) *Efecto reactivo o interactivo del pretest.* La aplicación de un test previo puede restringir el alcance de la generalización de los resultados obtenidos. El pretest puede entrenar a los sujetos en relación con la variable independiente; así, los resultados obtenidos no serán generalizables a un grupo de sujetos que no hayan pasado el pretest.

d) *Inteiferendas de tratamientos múltiples.* Cuando se administra una serie de tratamientos, a menos que el efecto de los tratamientos anteriores quede eliminado con el tiempo, no se sabrá si los sujetos reflejan el efecto del tratamiento que se les aplica en ese momento o la posible interacción existente entre dicho tratamiento y los efectos de los tratamientos anteriores.

e) *Novedad del tratamiento.* La curiosidad, motivación o novedad pueden influir «aumentando» el efecto de la variable independiente. Cuando se aplicara el mismo tratamiento en un contexto real perdería efectividad a medida que se prolongara el tratamiento.

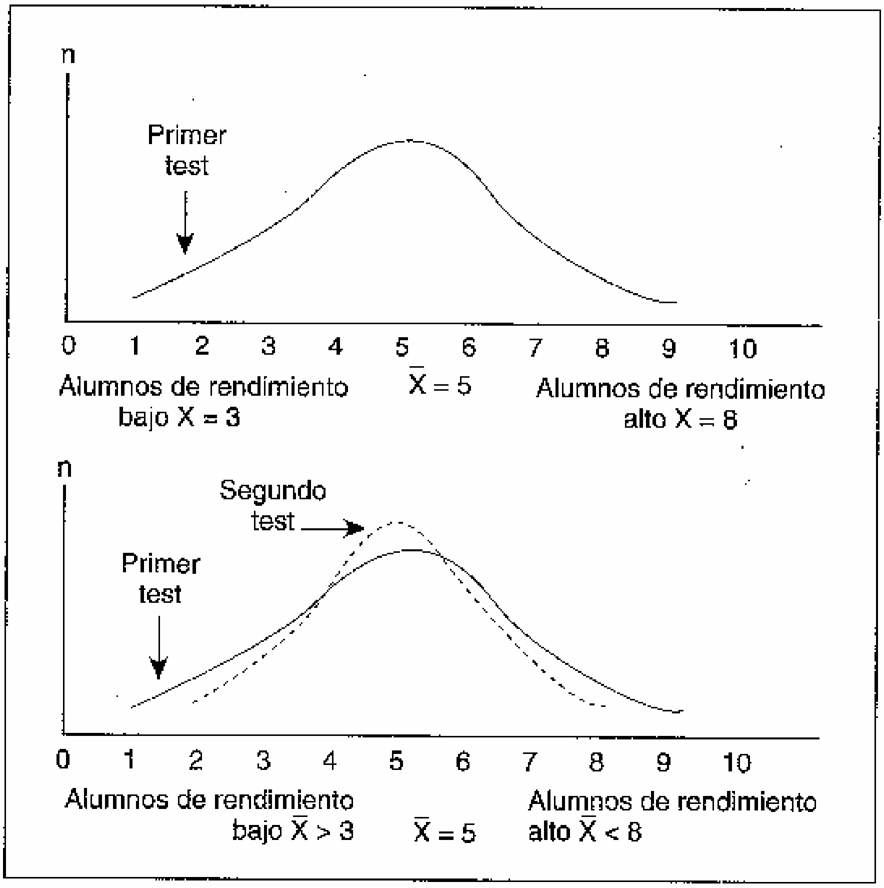


FIG. 5. 3 Efecto de la regresión estadística cuando se aplica un test por segunda vez en grupos extremos

5.5 CRITERIOS DE CLASIFICACIÓN DE LOS DISEÑOS

El método experimental puede adoptar diversas modalidades, siendo muchos los criterios de clasificación que pueden utilizarse: técnicas de control (Amau, 1978), número de grupos (Jiménez, 1985), mayor o menor rigurosidad, validez interna y externa (Campbell y Stanley, 1979), número de variables de tratamiento (Kirk, 1982). La adopción de uno u otro tipo de criterio genera taxonomías diferenciadas.

Didácticamente, el criterio del control parece ser el que inicialmente ofrece más ventajas, al desglosar los diseños en tres categorías: grupos al azar, bloques homogéneos e intragrupo. Dentro de cada categoría pueden efectuarse subdivisiones según el criterio del número de variables independientes y del número de niveles de tratamiento; así ocurre con los experimentos bivalentes, multivalentes y factoriales.

En la figura 5.4 pueden verse los diseños más importantes con asignación intergrupo o intragrupo y clasificados según el número de variables y las técnicas de control.

Para elegir el diseño que mejor se adapte a las necesidades del investigador y del problema planteado hay que tener en cuenta criterios como los siguientes:

1) Número de variables independientes que intervienen en la investigación. Cuando interviene una sola variable independiente hay que optar por los diseños unifactoriales o simples, y si están implicadas más de una variable independiente hay que elegir entre los multifactoriales o simplemente denominados factoriales.

2) Número de variables dependientes. Nuestra exposición se centrará siempre en diseños que contemplen una sola variable dependiente.

3) La asignación de los sujetos a los niveles de la variable independiente. Podemos elegir entre la asignación intergrupo e intragrupo. Si se opta por la asignación intergrupo podemos formar los grupos completamente al azar o según bloques. Como puede apreciarse en la figura 5.4, las diferencias individuales serán máximas en los diseños completamente al azar y mínimas en los diseños intragrupo.

4) Número de niveles de la variable independiente. Cuando dicha variable tiene dos niveles se utiliza un diseño bivalente, y si tiene más de dos se utiliza un diseño multivalente.

En los apartados siguientes se describen brevemente los diseños más utilizados.

5.6 DISEÑOS COMPLETAMENTE AL AZAR

Después de elegir la muestra al azar, en estos diseños los grupos se forman asignando al azar cada sujeto a uno de los grupos del experimento, y posteriormente, asignando también aleatoriamente cada nivel de tratamiento a los distintos grupos. Para que el azar pueda garantizar la equivalencia de los grupos hay que recurrir a muestras suficientemente grandes. De esta manera puede esperarse que las posibles variables contaminadoras se compensen y afecten por igual a todos los grupos antes de empezar los tratamientos.

Dentro de esta modalidad los diseños más utilizados son los bivalentes, multivalentes y factoriales.

Disminuyen las diferencias individuales →
Se reduce el error experimental

CRITERIOS / DISEÑOS	1.º Núm. VI	2.º Núm. VD	3.º ASIGNACIÓN EXPERIMENTAL — INTERGRUPOS: Completamente al azar	INTERGRUPOS: De bloques	INTRAGRUPOS
UNIFACTORIAL O SIMPLE	1	1	* Bivalente: Pretest - postest, Sólo postest * Multivalente: Pretest- postest, Sólo postest	* Grupos apareados * Bloqueo simple	* Tratamientos por sujetos * Análisis de tendecias
MULTIFACTORIAL O FACTORIAL	> 1	1	* Factorial	* Factorial bloques al azar	* Factorial intragrupo
* Análisis de la covarianza					

FIG. 5.4 Clasificación de los diseños más importantes. (Tomado de M. Bartolomé, 1980)

DISEÑOS BIV ALENTES

Se utilizan cuando aplicamos dos niveles de la variable independiente y generalmente se componen de un grupo experimental y un grupo control, o bien de dos grupos experimentales (sin grupo control). Cuando se utiliza el grupo control es posible equilibrar la influencia de las variables extrañas, pues tenderán a influir con la misma intensidad en ambos grupos. De esta manera muchas variables extrañas pueden quedar neutralizadas.

Supongamos que pretendemos detectar el efecto relativo del tiempo de estudio (A) sobre la asimilación de un contenido (VD). Si queremos analizar el efecto de 2 horas de estudio (a), y pasamos una prueba previa para comprobar la homogeneidad o equivalencia inicial y conocer el nivel de la variable dependiente antes de aplicar la independiente, utilizaremos un diseño pretest-postest (tabla 5.16) con grupo control. A este grupo le asignaremos el nivel mínimo de la variable independiente.

El grupo control es un grupo de referencia, pues, al igual que en el experimental, las mismas variables extrañas contextuales influirán con la misma intensidad en ambos grupos. En consecuencia, las diferencias observadas al medir la variable dependiente en cada grupo podrán atribuirse al tratamiento aplicado, siempre que los grupos sean inicialmente equivalentes. El grupo control permite estimar los cambios debidos a la maduración $\bar{X}_1 - \bar{X}_2$ en aquellos experimentos suficientemente prolongados.

TABLA 5.16 Diseño bivalente pretest - postest

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	$\bar{X}_1$	a_0	$\bar{X}_2$
2 Experimental	25	Azar	$\bar{X}_3$	a_1	$\bar{X}_4$

El pretest puede aplicarse a la variable dependiente o a una variable extraña con la que puede estar relacionada. Si omitimos la primera prueba y asumimos la equivalencia inicial de los grupos, el diseño se denomina *sólo postest* (tabla 5.17). Con esta modalidad podemos comparar los niveles alcanzados por la variable dependiente en el grupo control y en el experimental, pero no podemos estimar los cambios operados en el grupo control y en el experimental pomo poder comparar los niveles antes y después del tratamiento.

TABLA 5.17 Diseño bivalente sólo postest

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	-	a_0	$\bar{X}_1$
2 Experimental	25	Azar	-	a_1	$\bar{X}_2$

DISEÑOS MULTIV ALENTES

Se utiliza cuando se aplican tres o más niveles de la variable independiente (tabla 5.18).

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	$\bar{X}_1$	a_0 (oh)	$\bar{X}_2$
2 Experimental	25	Azar	$\bar{X}_3$	a_1 (2h)	$\bar{X}_4$
3 Experimental	20	Azar	$\bar{X}_5$	a_2 (3h)	$\bar{X}_2$

Tanto en diseños bivalentes como multivalentes (pretest-postest), si se demuestra la equivalencia inicial (homogeneidad) de los grupos, bastará comparar las medias del

postest para comprobar el efecto relativo de las horas de estudio en el rendimiento (tabla 5.19).

TABLA 5.19 Posibles comparaciones que pueden efectuarse en diseños bivalentes y multivalentes

Objetivo	Comparación	Prueba
Homogeneidad o equivalencia inicial	$\bar{X}_1 - X_3 - X_5$	ANOVA
Homogeneidad final (prueba hipótesis, efecto VI)	$\bar{X}_2 \quad \bar{X}_4 \cdot \bar{X}_6$	ANOVA
Cambio en grupos experimentales	$\bar{X}_4 - \bar{X}_3$	t apareados
	$X_6 \cdot \bar{X}_5$	t apareados
Cambio en grupo de control (Influencia externa, madurez)	$\bar{X}_2 - \bar{X}_1$	t apareados

Si los grupos experimentales muestran un mayor cambio que el control puede inferirse que el tratamiento origina la diferencia.

Siempre que sea posible, es más adecuado utilizar diseños multivalentes, ya que los diseños bivalentes ofrecen una generalización muy restringida de las conclusiones que no serán extensibles a otros niveles de VI.

DISEÑOS FACTORIALES COMPLETAMENTE AL AZAR

Los diseños bivalentes y multivalentes constan sólo de una variable independiente con diferentes niveles, cada uno de los cuales se aplica a distinto grupo de sujetos. Dado que el investigador sólo está interesado en la acción de una variable independiente, se denominan *diseños simples* o *diseños unifactoriales.* En estos diseños se intenta mantener todo constante excepto la variable independiente, con el fin de poder atribuir a dicha variable los cambios observados en la variable dependiente. Sin embargo, se plantean dos dificultades:

1) Es casi imposible mantener constantes todas las variables excepto una. El control completo es muy difícil de conseguir.

2) En la realidad educativa es apropiado hablar de causación múltiple, pues en cada experimento una gran multiplicidad de variables afectan conjuntamente a la variable dependiente.

En consecuencia, más que tratar de hacer lo imposible para mantener constantes todas las variables extrañas (situación artificial y poco generalizable), podemos considerar la acción conjunta de dos o más variables independientes.

En un diseño multifactorial o «factorial», el investigador manipula dos o más variables independientes (cada una de ellas con dos o más niveles de tratamiento) con el fin de estudiar los cambios operados en la variable dependiente como consecuencia de dicha manipulación.

En los diseños factoriales se establecen todas las posibles combinaciones de los niveles de una variable independiente con los niveles de otra o más variables independientes. A cada combinación o casilla se asigna, al azar, un grupo de sujetos. Al diseño factorial de dos variables o factores se le denomina *bifactorial,* y si cada uno de ellos tiene dos niveles, se expresa mediante la notación 2 x 2, y exige formar 2 x 2 = 4 grupos experimentales o casillas. Un diseño bifactorial con tres niveles en cada factor se denomina 3 x 3. Así, un diseño 2 x 3 indica un experimento con dos factores, el primero con dos niveles y el segundo con tres. De la misma manera, un diseño 3 x 2 x 4 se refiere a la acción conjunta de tres factores o variables, la primera de las cuales actúa con tres niveles de tratamiento, la segunda con dos y la tercera con cuatro. Para aplicar este último diseño harían falta 3 x 2 x 4 = 24 grupos o casillas.

Aunque es casi ilimitado el número de variables que se pueden manipular, nos centraremos en diseños bifactoriales a dos niveles de acción (2 x 2), es decir, dos factores o variables independientes con dos niveles cada una. A partir de estos diseños se puede generalizar a otro tipo de diseños más complejos.

Veamos el siguiente ejemplo. Con 16 sujetos elegidos al azar se pretende comprobar cómo influyen en el rendimiento el método didáctico utilizado (pasivo o activo según el grado de participación del alumno) y la gratificación otorgada (alta o baja según la mayor o menor frecuencia de recompensas: bolígrafos, libros, pinturas, cromos, etc.) cada vez que el alumno asimila correctamente un contenido temático determinado. Podría establecerse una hipótesis como la siguiente: el tipo de método tiene un efecto diferente según se apliquen gratificaciones bajas o altas. Con el fin de facilitar la recogida de datos puede esquematizarse la matriz factorial (2 x 2) donde cada letra mayúscula representa una variable o factor, y las respectivas minúsculas, los niveles correspondientes. Los subíndices indican la magnitud de cada nivel del factor implicado, es decir, $a_1 < a_2$, $b_1 < b_2$, 10 que da lugar a los siguientes símbolos:

		A	
		a_1	a_2
B	b_1	$a_1 b_1$	$a_2 b_1$
	b_2	$a_1 b_2$	$a_2 b_2$

Para llevar a cabo el experimento habría que formar cuatro grupos de sujetos al azar y asignarles una combinación de tratamientos también al azar. La representación simbólica del diseño 2 x 2 planteado también puede hacerse del siguiente modo:

Grupos	*n*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	4	Azar	-	a_1b_1	$\bar{X}_1$
2	4	Azar	-	a_1b_2	$\bar{X}_2$
3	4	Azar	-	a_2b_1	$\bar{X}_3$
4	4	Azar	-	a_2b_2	$\bar{X}_4$

Las respectivas medias de cada grupo se representarían del siguiente modo:

A
Método

B
Gratificación

	Pasivo	Activo
Baja	$\bar{X}_{,}$	$\bar{X}_3$
Alta	$\bar{X}_2$	$\bar{X}_4$

Las combinaciones de tratamientos, tal como han sido simbolizadas más arriba, son las siguientes:

A
Método

B
Gratificación

	Pasivo	Activo
Baja	Grupo 1	Grupo 3
Alta	Grupo 2	Grupo 4

Grupo 1: Las dos variables actúan con los niveles más bajos, es decir, método pasivo o basado en una participación mínima del alumno y en la poca adjudicación de materiales escolares como recompensa (gratificación baja).
Grupo 2: Los sujetos se someten al nivel más bajo de participación (método pasivo) mientras que la gratificación es máxima.
Grupo 3: Se concede a los alumnos el mayor grado de iniciativa y participación en las clases (método activo), pero el nivel de gratificación es el más bajo.
Grupo 4: Se combinan los niveles más altos de cada factor o variable: método activo y gratificación alta.

EFECTOS FACTORIALES

Los diseños factoriales aportan una información que no podría obtenerse en un diseño unifactorial de una sola variable independiente aunque ésta tenga más de dos niveles de tratamiento. Con un diseño factorial pueden analizarse los siguientes efectos:

a) *Efectos simples.* Expresan los efectos de los niveles de una variable independiente o factor dentro de cada nivel del otro factor. Así, en el ejemplo propuesto, el factor A (método) ejerce dos efectos simples. El primer efecto simple de A se obtiene al comparar las medias de A en el primer nivel del factor B (gratificación baja): X3-X¡, El segundo efecto simple de A se obtiene al comparar las medias de A en el segundo nivel del factor B (gratificación alta): X_4-X_2. De la misma manera se pueden obtener los efectos simples del factor B (gratificación).

b) *Efectos principales o globales.* Se refieren a la influencia global de cada variable independiente por separado, es decir, sin tener en cuenta los valores que toma la otra variable independiente. El efecto principal equivale a la influencia ejercida cuando en el experimento interviene una sola variable independiente. En consecuencia, los efectos principales se pueden estimar calculando las diferencias entre las medias de las filas y entre las medias de las columnas. Ahora bien, dado que los efectos simples son estimaciones del efecto principal, para calcular el efecto principal también pueden promediarse los efectos simples.

c) *Efectos secundarios o de interacción.* Existe interacción entre dos variables independientes cuando el valor que toma la variable dependiente, como resultado de las modificaciones introducidas en aquéllas, depende de los valores que toma la otra variable independiente. La interacción, por tanto, supone que un tratamiento tiene diferentes efectos al combinado con diferentes niveles de otro tratamiento. En definitiva, la influencia simultánea de varios factores hace que ciertas combinaciones entre ellos sean más efectivas que otras.

En el ejemplo propuesto, gracias al diseño factorial podemos comprobar si el efecto del método activo (o del método pasivo) varía a medida que se aplica combinado con los respectivos niveles de la variable gratificación (alta y baja). Si la variación fuera efectiva (significativa) es que existe interacción entre las dos variables o factores (método y gratificación).

La interacción es significativa cuando la diferencia entre los efectos simples de un factor es demasiado grande para ser atribuible al azar. Así, la comparación de los efectos simples para un determinado factor es una estimación de la posible existencia de interacción. En un diseño factorial 2 x 2 la interacción también puede estimarse calculando la diferencia entre la suma de los valores de las diagonales en la tabla de datos promedio.

Para determinar si existe interacción entre dos o más variables puede procederse de dos maneras:

1) *Análisis de la varianza.* Proporciona información sobre la significación estadística de la interacción. Su estudio pormenorizado se expondrá en capítulos posteriores al tratar el análisis de datos.

2) Representación gráfica. La inspección visual de la gráfica permite apreciar de forma intuitiva si existe o no interacción. En el eje de abscisas se representan los valores de uno de los factores y en el de ordenadas se representan los valores de la variable dependiente. Al unir los puntos correspondientes a los niveles del segundo factor, las rectas resultantes pueden tomar alguna de las siguientes formas:

- rectas paralelas;
- rectas que no son paralelas, aunque no llegan a cortarse en la gráfica;
- rectas que se cortan en la gráfica.

Ilustraremos la interpretación gráfica de la interacción a partir del ejemplo ya planteado para comprobar cómo influyen en el rendimiento el método didáctico utilizado (pasivo-activo) y la gratificación (alta-baja). También se describirán las tres situaciones contemplando distintos tipos de efectos factoriales como son: los efectos principales o globales, secundarios o de interacción, y simples.

PRIMERA SITUACIÓN: RECTAS PARALELAS

Supongamos que en el ejemplo planteado las puntuaciones de rendimiento obtenidas por los 16 sujetos en el postest fueran las siguientes:

Sujeto	*Tratamiento*	*Rendimiento*
1	a_1b_1	0
2	a_1b_1	1
3	a_1b_1	2
4	a_1b_1	1
5	a_1b_2	3
6	a_1b_2	4
7	a_1b_2	3
8	a_1b_2	2

Sujeto	*Tratamiento*	*Rendimiento*
9	a_2b_1	4
10	a_2b_1	5
11	a_2b_1	6
12	a_2b_1	5
13	a_2b_2	8
14	a_2b_2	6
15	a_2b_2	8
16	a_2b_2	6

Estos datos pueden tabularse de distintas maneras. Cuando queremos analizar los efectos globales de cada variable puede ser útil una tabla como la siguiente:

		A		
		a_1	a_2	$\bar{X}$
B	b_1	0 1 2 1	4 5 6 5	3
	b_2	3 4 3 2	8 6 8 6	5
	$\bar{X}$	2	6	4

Como ejemplo ilustrativo se han registrado todas las puntuaciones directas de rendimiento; sin embargo, presenta mayor utilidad incluir las medias obtenidas por cada grupo en una tabla de datos promedio como se verá más adelante.

Para apreciar con más claridad el efecto del método en cada nivel de gratificación también puede elaborarse una tabla como la siguiente:

Método	*Pasivo*		*Activo*	
Gratificación	Baja	Alta	Baja	Alta
Puntuaciones	1	4	5	6
de rendimiento	2	3	6	8
	1	2	5	6
Rendimiento medio	1	3	5	7

A partir de las tablas anteriores pueden estimarse los efectos factoriales principales o globales, secundarios o interacciones, y simples o diferenciadores.

a) *Efectos factoriales principales o globales*

Se refieren a la influencia global de las variables método y gratificación consideradas por separado. En definitiva, es como si hubiéramos realizado dos experimentos bivalentes, uno para ver qué método es más efectivo (el activo o el pasivo) y otro para ver qué gratificación es más efectiva (la alta o la baja). La comparación de los resultados finales de cada columna nos permite comprobar si los sujetos rinden más con un método pasivo que con un método activo. Al comparar los resultados de las filas podemos estimar la eficacia relativa de cada tipo de gratificación (baja o alta):

1) Efecto global de la gratificación. El efecto principal de la gratificación vale 5 - 3 = 2. Es decir, al pasar de un nivel de gratificación bajo a un nivel alto el rendimiento aumenta en 2 unidades:

Gratificación	*Baja*	*Alta*
Puntuaciones de rendimiento	0	3
	1	4
	2	3
	1	2
	4	8
	5	6
	6	8
	5	6
Rendimiento medio	3	5

2) Efecto global del método. El efecto principal del método vale 6- 2 = 4. Es decir, al pasar de un método pasivo a un método activo el rendimiento aumenta en cuatro unidades:

Método	*Pasivo*	*Activo*
Puntuaciones de rendimiento	0	4
	1	5
	2	6
	1	5
	3	8
	4	6
	3	8
	2	6
Rendimiento medio	2	6

b) *Efecto factorial secundario o interacción*

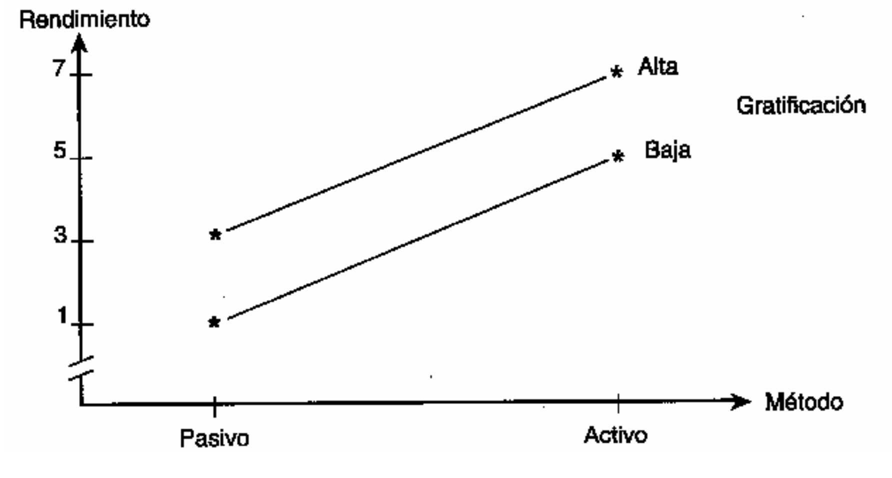
Las conclusiones anteriores de que la gratificación alta es más efectiva que la baja porque aumenta el rendimiento en dos unidades (suponiendo que hubiera significación estadística) y de que el método activo es más efectivo que el pasivo tienen sentido en el caso de que no existan efectos factoriales secundarios o interacciones.

No existe interacción entre la gratificación y el método cuando el aumento del rendimiento provocado por la gratificación alta (dos unidades) es el mismo al combinarse con el método pasivo y con el activo. Es decir, al pasar de una gratificación baja a otra alta el rendimiento aumenta una cantidad constante que siempre vale dos unidades, tanto si el método es pasivo como si es activo. Esto significa que los efectos simples correspondientes a los niveles de una misma variable o factor son estadísticamente iguales. La tabla de datos promedio es la siguiente:

Método

Gratificación	Pasivo	Activo	
Baja	1	5	3
Alta	3	7	5
	2	6	4

Puede comprobarse que al pasar de un método pasivo a un método activo el rendimiento también aumenta una cantidad constante de cuatro unidades, tanto para la gratificación baja como para la gratificación alta. Es decir, los efectos simples vuelven a ser iguales. La ausencia de interacción o igualdad de efectos simples queda reflejada por la presencia de dos rectas paralelas:



La interacción puede estimarse a partir de la diferencia entre la suma de los valores de las diagonales:

$$[(1 + 7) - (3 + 5)] /2 = O.$$

Cuando la interacción no es significativa se puede concluir que los efectos factoriales son «aditivos» o que existe aditividad, y por tanto, que existe independencia entre los factores o variables. Aunque la interacción no sea significativa, en algunos casos el análisis de los efectos simples puede ser muy revelador.

c) *Efectos factoriales simples*

Si nos fijamos en la variable gratificación, podemos determinar dos efectos de la misma:

1) Efecto simple de la gratificación cuando el método es pasivo: 3- 1 = 2.
2) Efecto simple de la gratificación cuando el método es activo: 7- 5 = 2.

Esta información indica que el efecto de la gratificación tiene la misma intensidad al combinarse con ambos métodos. Con respecto a la variable método puede procederse de la misma manera:

3) Efecto simple del método cuando la gratificación es alta: 7- 3 = 4.
4) Efecto simple del método cuando la gratificación es baja: 5- 1 = 4.

Estos resultados indican que el método tiene una influencia idéntica cuando la gratificación es baja y cuando es alta. Cuando las diferencias entre los efectos factoriales simples de una variable pueden atribuirse estadísticamente al azar no existen efectos factoriales secundarios o interacción. Así, para la variable gratificación, sus efectos simples (2 y 2) son dos estimaciones de la influencia de dicha variable. Si son estadísticamente iguales, hay que concluir que la variable método no interviene en ellos y que no existe interacción. Es decir, la variable gratificación, al combinarse con los dos tipos de método, influye por igual en el rendimiento.

Puede comprobarse que el efecto principal o influencia global de cada variable puede estimarse también a partir de los efectos simples:

- Influencia global de la gratificación: promedio de los efectos simples (2 + 2) / 2 = 2. Como ya hemos visto, dicho efecto se puede calcular recurriendo a la diferencia entre las medias de las filas: 5- 3 = 2.
- Influencia global del método: promedio de los efectos simples (4 + 4) / 2 = 4.

El mismo efecto se puede calcular con la diferencia entre las medias de las columnas: 6-2 = 4.

También puede estimarse la interacción a partir de la diferencia entre los efectos simples, ya sea 2- 2 = 0, o bien, 4- 4 = 0.

SEGUNDA SITUACIÓN: LAS RECTAS NO SON PARALELAS PERO NO SE CORTAN EN LA GRÁFICA

También podríamos haber obtenido datos ficticios como los siguientes:

		Método		
		Pasivo	Activo	
Gratificación	Baja	3	1	2
	Alta	5	7	6
		4	4	4

Según la tabla, los tipos de efectos son:

a) *Efectos factoriales principales* o *globales*

1) Influencia global de la gratificación: 6 - 2 = 4. Estadísticamente podría existir diferencia significativa y ser más efectivo el nivel alto de gratificación.
2) Influencia global del método: 4- 4 = O. El método activo tiene la misma efectividad que el método pasivo, ambos tienen un efecto idéntico.

b) *Efecto factorial secundario* o *interacción*

Las conclusiones anteriores son globales y pueden inducir a error. No son válidas por igual cuando cada variable se combina con los niveles de la otra. Según los efectos globales de la gratificación, al pasar de un nivel bajo de gratificación a un nivel alto aparece un aumento en rendimiento de cuatro unidades. Si este aumento fuera estadísticamente igual (cuatro unidades) bajo los dos métodos (activo y pasivo), no existiría interacción y el efecto global sería válido. Sin embargo, bajo el método pasivo el aumento es de sólo dos unidades, pero bajo el método activo el aumento es de seis unidades, lo que origina una aumento global medio de (2 + 6) / 2 = 4 unidades, como ya se había calculado. Este aumento es debido al cambio introducido en los niveles de gratificación.

Si sólo nos basamos en los efectos globales del método, el error puede ser aún más grave. Según los efectos principales o globales, tanto el método pasivo como el activo tienen la misma efectividad (4 - 4 = O). Ahora bien, cuando la gratificación es baja, al pasar de un método pasivo a un método activo el rendimiento sufre un descenso de dos unidades (1- 3 = -2). En cambio, cuando la gratificación es alta, aparece un aumento de 7 - 5 = 2 unidades; de ahí que se detecte un aumento global medio que vale

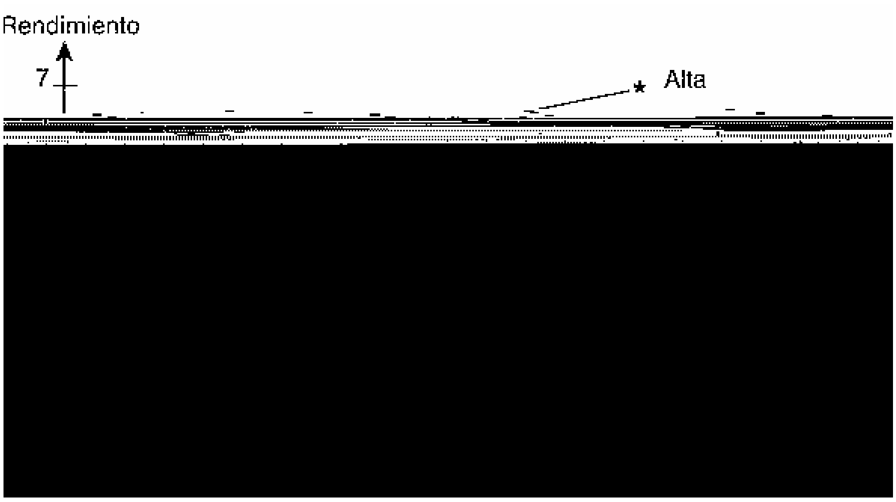
[2 + (- 2)] /2 = O.

Es decir, a pesar de que ambos métodos tienen el mismo efecto global, sus efectos factoriales simples son diferentes para cada nivel de gratificación.

Se produce una interacción entre las variables método y gratificación cuando los cambios observados en el rendimiento como consecuencia del método son significativamente diferentes para cada tipo de gratificación; es decir, la efectividad del método es distinta según el nivel de gratificación, y viceversa. El cálculo numérico de la interacción sería

$$[(3 + 7) - (5 + 1)] / 2 = 2$$

Al existir interacción no es suficiente analizar los efectos globales o principales de cada variable; es imprescindible analizar los efectos simples. Cuando existe interacción las rectas no tienden a ser paralelas:



c) *Efectos factoriales simples*

Con respecto a la variable gratificación hemos estimado dos efectos de la misma:

1) Efecto factorial simple de la gratificación cuando el método es pasivo: 5-3 = 2.
2) Efecto factorial simple de la gratificación cuando el método es activo: 7-1 = 6.

En consecuencia, el efecto de la gratificación tiene diferente intensidad (debería comprobarse si la diferencia es excesiva para ser atribuida sólo al azar) cuando el método es pasivo y cuando es activo.

Con respecto a la variable método puede procederse de la misma manera:

3) Efecto factorial simple del método cuando la gratificación es alta: 7-5 = 2.
4) Efecto factorial simple del método cuando la gratificación es baja: 1-3 = -2.

Como puede apreciarse, el efecto es negativo porque el nivel más participativo del método (activo) es menos efectivo que el nivel más bajo de participación (pasivo).

Estos resultados indican que el método tiene una influencia distinta según sea el nivel de gratificación (alta o baja). Cuando la gratificación es alta, al pasar de un método pasivo a un método activo se aprecia que el rendimiento aumenta en dos unidades. En cambio, cuando la gratificación es baja, ocurre al contrario: al pasar de un método pasivo a un método activo el rendimiento disminuye en dos unidades. Como es evidente, los efectos globales pueden calcularse recurriendo a los efectos globales:

- Influencia global de la gratificación: promedio de los efectos simples (2 + 6) *12* = 4. El mismo efecto se puede calcular a partir de la diferencia entre las medias de las filas: 6 - 2 = 4. - Influencia global del método: promedio de los efectos simples [2 + (- 2)] *12* = 0.

De la misma manera, a partir de la diferencia entre las medias de las columnas: 4 - 4=0.

TERCERA SITUACIÓN: LAS RECTAS SE CORTAN EN LA GRÁFICA

Otra posibilidad es que los datos resultantes sean como los de la tabla que se adjunta:

a) *Efectos factoriales principales o globales*

1) Influencia global de la gratificación: 5 - 5 = 0. Los dos niveles de gratificación tienen la misma efectividad.

2) Influencia global del método: 6- 4 = 2. El método activo aumenta dos unidades el rendimiento con respecto al método pasivo.

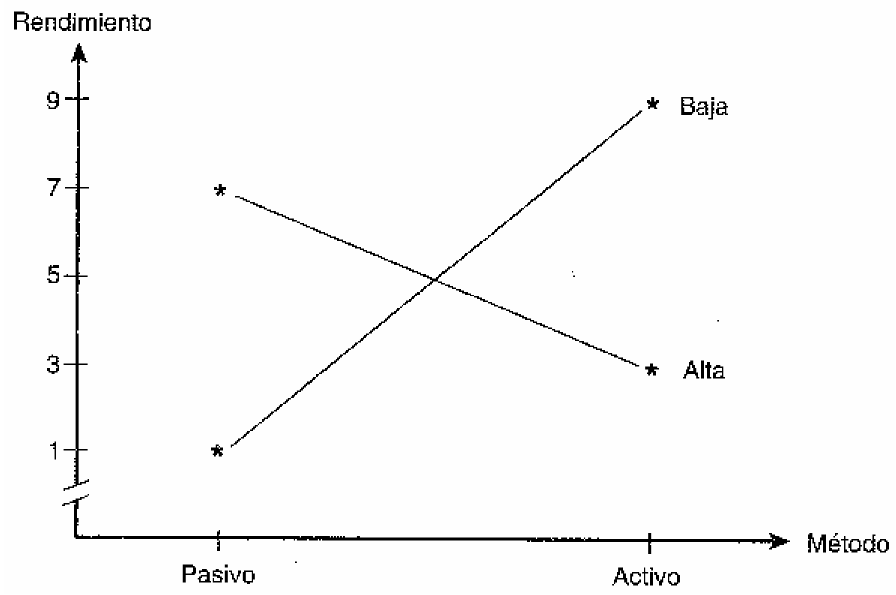
Método

		Pasivo	Activo	
Gratificación	Baja	1	9	5
	Alta	7	3	5
		4	6	5

b) *Efecto factorial secundario a interacción*

Cuando en la gráfica aparecen rectas que se cortan es muy posible que exista interacción entre los dos factores (gratificación y método), cuyo valor sería:

[(1 + 3) - (7 + 9)] *12* = - 6



Cuando exista interacción es necesario analizar los efectos simples.

c) *Efectos factoriales simples*

1) Efecto factorial simple de la gratificación cuando el método es pasivo: 7 - 1 = 6.
2) Efecto factorial simple de la gratificación cuando el método es activo: 3 - 9 = - 6.

Con respecto a la variable método puede procederse de la misma manera:

3) Efecto factorial simple del método cuando la gratificación es alta: 3 - 7 = - 4.
4) Efecto factorial simple del método cuando la gratificación es baja: 9 - 1 = 8.

Para comprobar los efectos globales:

- Influencia global de la gratificación: promedio de los efectos simples:

[6 + (- 6)] / 2 = 0, o bien 5 - 5 = 0.

- Efecto global del método: [8 + (- 4)] /2 = 2, o bien, 6 - 4 = 2.

5.7 DISEÑOS DE BLOQUES HOMOGÉNEOS AL AZAR

En los diseños completamente al azar la heterogeneidad entre los sujetos puede originar una varianza intragrupo (o del error) muy elevada. Con el fin de reducir la variabilidad de los sujetos y hacer más sensible el experimento se puede introducir la

técnica del bloqueo. De esta manera se pueden apreciar mejor los efectos de VI. Supongamos que estamos interesados en ver el efecto de tres niveles de gratificación (A = alto, M = medio y B = bajo) en el rendimiento de 12 niños de primaria. Para realizar un experimento de bloques hay que tener en cuenta las siguientes fases:

1) Mídase alguna variable extraña que sospechemos que pueda estar estrechamente relacionada con la variable dependiente. En consecuencia, podría contaminar el experimento. Dicha variable se denomina *variable de bloqueo* o, en su caso, *de apareo* (cuando cada bloque está formado por dos sujetos). Generalmente, dicha variable no coincide con la variable dependiente. En el ejemplo planteado consideramos que la variable inteligencia sería una variable de bloqueo adecuada. Después de medir la inteligencia obtenemos las siguientes puntuaciones:

Sujetos	*Inteligencia*
Antonio	98
Rosa	100
Carlos	94
Ignacio	94
Gema	96
Maribel	98

Sujetos	*Inteligencia*
Mónica	96
José	100
Laura	94
Ana	98
Nuria	100
Raquel	96

2) En función de las puntuaciones obtenidas en la variable de bloqueo, fórmense grupos de sujetos (bloques) con puntuaciones iguales o lo más similares posible. Dentro de cada bloque se asignan al azar, como mínimo, tantos sujetos o unidades experimentales como niveles tenga la variable independiente:

Bloques	*CI*		*Sujetos*	
Bloque 1	94	Ignacio	Carlos	Laura
Bloque 2	96	Gema	Mónica	Raquel
Bloque 3	98	Maribel	Antonio	Ana
Bloque 4	100	Nuria	José	Rosa

3) Asígnese cada nivel de la variable independiente al azar:

			Gratificación	
Bloques	*CI*	*Baja*	*Media*	*Alta*
Bloque 1	94	Ignacio	Carlos	Laura
Bloque 2	96	Gema	Mónica	Raquel
Bloque 3	98	Maribel	Antonio	Ana
Bloque 4	100	Nuria	José	Rosa

4) Aplíquense los niveles de la variable independiente.
5) Mídase la variable dependiente:

		Gratificación		
Bloques	el	*Baja*	*Media*	*Alta*
Bloque 1	94	9	9	9
Bloque 2	96	8	7	6
Bloque 3	98	6	5	4
Bloque 4	100	4	4	2

Si cada bloque está formado por dos sujetos, puede tratarse de un diseño de grupos apareados. Cuando cada unidad experimental está formada por grupos de sujetos, la variable de bloqueo puede considerarse como otra variable independiente o factor, y el experimento se convierte en factorial.

El diseño planteado se denomina *de bloques al azar* 3 4 porque el factor manipulado adopta tres niveles y se han formado cuatro bloques de sujetos con la variable de bloqueo. La representación simbólica del diseño sería la siguiente:

Grupos	*n*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	3	BA_4	-	a_1	$\bar{Y}_1$
2	3	BA_4	-	a_2	$\bar{Y}_2$
3	3	BA_4	-	a_3	$\bar{Y}_3$
4	3	BA_4	-	a_4	$\bar{Y}_4$

La técnica del bloqueo adolece de tres limitaciones importantes. En primer lugar, cuando se ha elegido un sujeto o grupo de sujetos para formar un bloque, el investigador debe seleccionar más sujetos semejantes entre los restantes para completar el bloque iniciado. Con frecuencia es difícil encontrar sujetos suficientes. Esta circunstancia origina la segunda limitación: los sujetos que no son similares a los incluidos en los bloques deben excluirse, con lo que disminuye la representatividad de la muestra con respecto a la población. Por último, la homogeneidad entre los sujetos de un mismo bloque se establece con respecto a la variable de bloqueo, pero los sujetos incluidos en un mismo bloque pueden diferir en alguna otra variable.

5.8 DISEÑOS INTRAGRUPO O DE MEDIDAS REPETIDAS

La eficacia y sensibilidad de un experimento aumenta en proporción a la homogeneidad de los sujetos, ya que, con grupos experimentales más homogéneos, se tiene mayor probabilidad de comprobar, en el caso de que exista, la posible influencia

de la variable independiente sobre la variable dependiente. La varianza del error disminuye progresivamente al pasar de los diseños completamente al azar a los diseños de bloques homogéneos. Con los diseños intragrupo la técnica del bloqueo se lleva hasta sus últimas consecuencias. Si a cada sujeto se le aplican todos los niveles de la variable independiente y se comparan entre sí los efectos producidos, hemos llevado hasta el extremo las condiciones de homogeneidad con respecto a las diferencias individuales. De hecho, cada sujeto puede considerarse como un solo bloque.

En los diseños intragrupo es importante el orden de aplicación. Los resultados obtenidos en la variable dependiente pueden depender de la secuencia seguida en la aplicación de los tratamientos. Para evitarlo es viable aleatorizar el orden de aplicación o puede procederse a un contrabalanceo de los tratamientos.

La representación simbólica de un diseño intragrupo puede realizarse del siguiente modo:

Grupos	*número de sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n	A	-	$a_1; a_2; a_3$	$\bar{X}_1; \bar{X}_2; \bar{X}_3$

Veremos dos aplicaciones de los diseños intragrupo: el diseño tratamiento por sujetos y el análisis de tendencias.

DISEÑO DE TRATAMIENTOS POR SUJETOS

Cuando estamos interesados en comparar la efectividad de una serie de tratamientos o niveles de la variable independiente se utiliza el denominado *diseño de tratamientos por sujetos.*

La representación simbólica anterior puede hacerse extensible del siguiente modo:

Grupos	*Pretest (VI)*	*Tratamiento (VD)*	*Postest*
1	-	a_1 a_2 $a_3 \dots a_k$	X_1 X_2 $X_3 \dots X_k$
1	-	a_1 a_2 $a_3 \dots a_k$	X_1 X_2 $X_3 \dots X_k$
1	-	a_1 a_2 $a_3 \dots a_k$	X_1 X_2 $X_3 \dots X_k$
•			
•			
•			
n	-	a_1 a_2 $a_3 \dots a_k$	X_1 X_2 $X_3 \dots X_k$
			$\bar{X}_1$ $\bar{X}_2$ $\bar{X}_3 \dots \bar{X}_k$

En la medida de lo posible, el orden de aplicación de los niveles de la variable independiente será totalmente aleatorio. La estructura lógica de este diseño es similar a la del diseño de bloques, ya que cada sujeto equivale a un bloque.

ANÁLISIS DE TENDENCIAS

Se utiliza cuando el énfasis se centra en la mejora o cambio que se produce como consecuencia de una práctica más o menos prolongada. Los niveles de la variable independiente difieren entre sí según una cantidad constante. Se trata de comprobar si las medias obtenidas en la variable dependiente bajo cada tratamiento tienden a variar en un sentido lineal, curvilíneo, etc., y hasta qué punto es significativa esta tendencia. Es un diseño muy común en situaciones de aprendizaje donde interesan los efectos acumulados de los tratamientos (fig. 5.5).

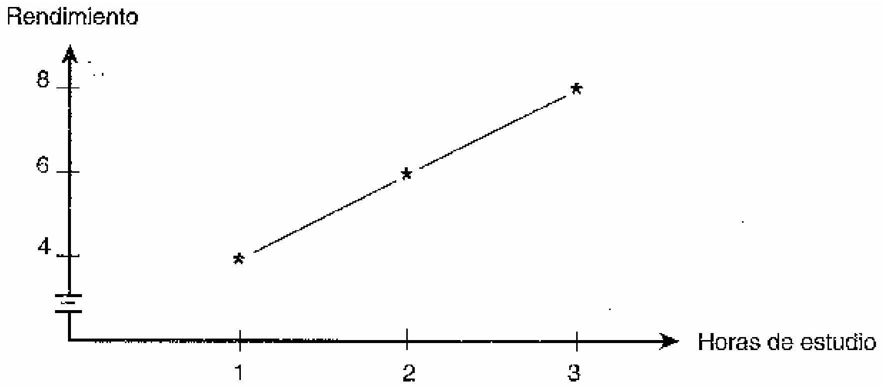


FIG. 5.5 Función lineal de las medias de rendimiento según las horas de estudio en un experimento intragrupo

5.9 TIPOS DE EXPERIMENTOS

Para conseguir el máximo grado de control posible en los experimentos se tiende a un aislamiento del contexto natural donde se produce el fenómeno. La finalidad de este aislamiento radica en el objetivo de eliminar las múltiples influencias extrañas capaces de actuar juntamente con la variable independiente, pues la acción de dichas variables la confundiríamos con la influencia de la variable independiente sobre la variable dependiente.

En la medida que controlamos las variables extrañas (restricciones) garantizamos la validez interna, pero las condiciones en que se desarrolla el experimento son más artificiales y, por tanto, menos representativas, pues en un contexto real las variables no se hallan tan controladas. En consecuencia, la generalización a situaciones reales queda reducida. De ahí la conveniencia de diferenciar entre experimentos de laboratorio y de campo, teniendo en cuenta que entre ambas modalidades sólo existe una diferencia de grado, los primeros con un control máximo, y los segundos, con un control menor.

Experimento de laboratorio. Es un estudio de investigación donde se consigue un aislamiento del contexto natural en que se produce el fenómeno con el fin de eliminar las múltiples influencias extrañas capaces de afectar a la variable dependiente.

Experimento de campo. Es un estudio de investigación en una situación real donde una o más variables independientes son manipuladas por el experimentador en condiciones controladas con el máximo rigor que permita la situación.

5.10 POSIBILIDADES Y LÍMITES

El método experimental se aplica a través de diseños específicos que suponen adoptar variaciones respecto a la categoría global, y que posibilitan responder de forma más acorde a determinados problemas y situaciones de investigación. En cuanto a sus aplicaciones a cuestiones o problemas educativos, los diseños más rigurosos se han utilizado principalmente en el área de tecnología educativa y didáctica, configurando la inmensa mayoría de trabajos que se han publicado.

a) *Posibilidades.* Según algunos autores (Mayntz, 1983), es el más prominente de todos los métodos de investigación en cuanto que es el único que permite constatar estrictamente relaciones causales. Al manipular o provocar el fenómeno, el investigador está en condiciones idóneas para inferir proposiciones causales (Cook y Campbell, 1979; Arnau, 1978; Haimson y Elfebein, 1985). Otras funciones básicas del método experimental son descubrir relaciones en condiciones puras y no contaminadas, probar predicciones derivadas de la teoría o de otra investigación y desarrollar teorías e hipótesis para elaborar sistemas teóricos.

Estas funciones son viables gracias al elevado grado de control que se ejerce sobre:

- Las variables independientes, mediante su modificación en cantidad, grado o cualidad.
- Las variables extrañas, cuyo dominio o control facilitará la maximización de la varianza primaria.
- La medición, que deberá ser fiable, precisa y válida.

El control relativamente completo permite una gran especificidad en las definiciones operativas de las variables, con lo que se favorece la precisión y la posibilidad de llevar a cabo las réplicas necesarias. La investigación queda muy bien delimitada, es inequívoca y los errores de medida son mínimos, ya que las condiciones del experimento se especifican minuciosamente. Como consecuencia, la validez interna es alta, con lo que estamos ante un enfoque adecuado para probar diferentes aspectos de la teoría.

Otra característica vinculada a la necesidad de inferir efectos de los tratamientos es la utilización de la comparación. Existen muchos tipos de comparación en la experimentación que responden a diferentes propósitos y originan distintas modalidades de diseños de investigación. También cabe destacar su precisión y sensibilidad a la hora de maximizar la varianza experimental y minimizar la del error, como consecuencia del control de las situaciones en las que se produce el experimento (Kerlinger, 1982). En resumen, las ventajas y posibilidades de esta metodología se pueden resumir en precisión y economía (Mouly, 1978).

Los *experimentos de laboratorio* son especialmente idóneos para cubrir objetivos de investigación como los siguientes:

- Constatar relaciones en condiciones «puras» e incontaminadas.
- Someter a prueba las predicciones derivadas de la teoría (hipótesis).
- Generar nuevas investigaciones.
- Refinar las teorías y las hipótesis. Ayudan a construir sistemas teóricos.

El experimento de laboratorio presenta ventajas como las siguientes:

- Control relativamente completo.
- Permite la asignación aleatoria y la manipulación de una o más variables independientes.
- Alto grado de especificidad en las definiciones operativas de las variables.
- Precisión. Bien delimitado e inequívoco. La varianza del error es pequeña, ya que las condiciones del experimento se especifican minuciosamente disminuyendo los errores.
- Alta validez interna.
- Flexibilidad: hay muchas posibilidades experimentales y se pueden poner a prueba variados aspectos de la teoría.
- Replicabilidad: posibilidad de repetir el experimento introduciendo pequeñas variantes.

Las ventajas principales de los experimentos de campo pueden ser las siguientes:

- Muy utilizados en pedagogía, se llevan a cabo en aulas, colegios y otras instituciones y comunidades. Las variables actúan en su propio medio natural. - Al menos en teoría permiten manipular las VI y asignar sujetos y tratamientos al azar.
- Las variables suelen tener un efecto más potente que las de laboratorio. Cuanto más realista es la situación más potentes serán las variables.
- Adecuados para estudiar las influencias sociales complejas, los procesos y cambios sociales en situaciones de la vida real.
- Pueden utilizarse para someter a prueba las teorías y para solucionar problemas prácticos.
- Flexibilidad y aplicabilidad a una gran variedad de problemas.

b) *Límites.* Pero los experimentos no están exentos de algunos inconvenientes.

Como contrapartida existen limitaciones al aplicar esta metodología al campo pedagógico, como son la selección de las variables, la naturaleza re activa de los sujetos, posibles sesgo s introducidos por el experimentador, limitaciones procedentes del muestreo, problemas deontológicos (Pérez Juste, *1985a;* Mialaret, 1972), además de los de generalización como consecuencia de la artificialidad del control (Cook y Campbell, 1979).

Así, dada la artificialidad de la situación experimental, puede ocurrir que el influjo de la variable independiente sobre la dependiente sea débil, ya que la situación se ha provocado con propósitos especiales. Por ello, hay que tener en cuenta que los experimentos, especialmente los de laboratorio, presentan poca generalización. Esta circunstancia disminuye su validez externa, con el consiguiente riesgo de interpretar o extrapolar erróneamente los resultados.

Algunas de las limitaciones comentadas deben matizarse si se tiene en cuenta que los experimentos, más que generalizar, pretenden constatar relaciones de causalidad en condiciones puras y no contaminadas, sometiendo a prueba las predicciones derivadas de la teoría. De ahí que **un** cometido fundamental del enfoque experimental sea refinar las teorías a través del contraste de hipótesis, con lo que se favorece la construcción de sistemas teóricos.

Aunque parece ser el método idóneo para garantizar la validez interna, algunos autores (Harsé y Secord, 1976) la cuestionan a la hora de estudiar la conducta social, señalando limitaciones como las que impone la misma naturaleza de la situación, la operacionalización de conceptos, el desconocimiento de parámetros relevantes y el tipo de interacción peculiar de los experimentos.

Algunas limitaciones de los experimentos de laboratorio son las siguientes:

- Debilidad de las variables independientes, ya que se han creado las situaciones con propósitos especiales.
- Artificialidad de la situación experimental si no se tienen en cuenta los objetivos de los experimentos de laboratorio (relaciones de causación entre variables). - Baja validez externa, poca generalización, pero su objetivo primordial no es generalizar.

En cuanto a los experimentos de campo, pueden destacarse las siguientes limitaciones:

- Carecen de **un** control riguroso.
- Objeciones ante la manipulación de las variables que puedan afectar a la educación o al rendimiento del alumno.
- Falta de precisión.
- Requieren en cierta medida **un** experto en trabajo social.

CAPÍTULO 5

Metodología experimental

El proceso general de la investigación ya se ha descrito en el capítulo 3; sin embargo, cada método presenta variaciones que le permiten adaptarse a las exigencias de problemas específicos de investigación. En consecuencia, es apropiado apuntar algunas características inherentes al método experimental, abordando posteriormente aspectos como la varianza y las técnicas para controlarla, las fuentes de invalidez y la estrategia que hay que seguir en los diseños más importantes. A su vez, se contemplan las posibilidades y límites de dichos diseños y se ofrecen orientaciones para planificar un experimento.

Los estudios experimentales estrictos se caracterizan por presentar los rasgos siguientes:

1) El investigador manipula la variable independiente y le asigna niveles.

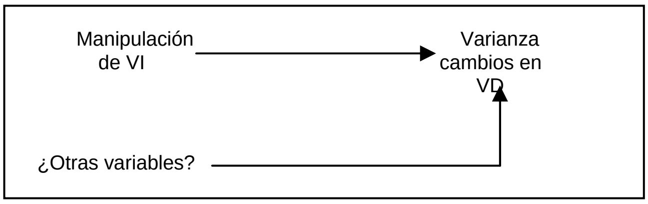
2) Han de aplicarse dos o más niveles de la variable independiente con el fin de poder contrastar los efectos sobre la variable dependiente. Estos efectos se estiman comparando los resultados obtenidos en la variable dependiente.

3) La muestra se elegirá al azar. Si los niveles de la variable independiente se aplican a grupos de sujetos distintos, antes de comenzar el experimento los sujetos se asignan al azar a los grupos donde se aplicarán los niveles de la variable independiente. A cada grupo se le aplica al azar un nivel de la variable independiente. Así se garantiza que por efecto del azar los grupos sean homogéneos o equivalentes, para que después de aplicar los niveles de la variable independiente dichos grupos sean comparables entre sí. De esta manera, las posibles diferencias encontradas al medir la variable dependiente en cada grupo son atribuibles a la aplicación de los niveles de la variable independiente.

5.1 TIPOS DE VARIANZA

La meta ideal del experimentador es que los cambios o variaciones en los valores de la variable dependiente (VD) sean atribuibles a las manipulaciones de la variable independiente (VI) y no a otras variables.

Como ya se ha expuesto, estas variables extrañas podrían influir con intensidad



desigual en los grupos control y experimental(es) y sus efectos se unirían al de la variable independiente, constituyendo fuentes de variación. En consecuencia, el investigador no podría distinguir qué cambios en la variable dependiente son exclusivamente debidos a la variable independiente, y el experimento quedaría invalidado, a menos que dichas variables se controlen igualando o eliminando su influencia. El objetivo del investigador es distinguir los cambios de la variable dependiente que se deben a la manipulación de la variable independiente y los que se deben a otras variables (extrañas). Tanto la variable independiente como las otras variables son fuentes de variación o varianza. Analizaremos las más importantes agrupadas en tres grandes bloques o fuentes de variación:

a) *El experimentador.* Las actitudes y expectativas del investigador sobre los efectos de los niveles de la variable independiente pueden influir en los resultados del experimento. También los rasgos físicos (edad, sexo, aspecto, etc.), la forma de comportarse y las características sociales (status, nivel de estudios, etc.) del experimentador pueden influir.

b) *El contexto experimental.* La posible contaminación de los datos en función del contexto puede provenir de dos fuentes: de cómo captan los sujetos las condiciones de la situación experimental y del diseño (nivel de las variables ambientales, condiciones de aplicación, etc). Veamos brevemente estos aspectos:

- *Demanda de la situación experimental.* La interpretación que los sujetos de los grupos experimentales realizan acerca del experimento puede condicionar su actuación. A través de esta interpretación el sujeto intuye lo que se espera de él (demanda de la situación experimental), y lo puede descubrir durante la realización del experimento, a través de la captación de indicios, y antes del experimento por medio de los rumores sobre la investigación, el lugar donde se realiza y quiénes lo llevan a cabo.

- *Relacionadas con el diseño.* Pueden provenir del medio ambiente (ruido, iluminación, temperatura, etc.) o de las condiciones de aplicación (instrucciones, sistemas de medición, presentación de materiales, estímulos y recogida de datos).

c) *El sujeto experimental.* Cuando los experimentos se llevan a cabo con sujetos universitarios, obligados o voluntarios, hay que cuestionar su representatividad con respecto al resto de la población. Antes de aplicarse los niveles de la variable independiente los grupos pueden diferir en variables como edad, sexo, motivación,

inteligencia o personalidad. Los sujetos de los grupos experimentales pueden verse más motivados por la novedad que los sujetos que pertenecen al grupo control.

Debido a que la acción de estos tres grupos de variables o fuentes de variación es simultánea, cuando se ha elegido una variable independiente, el investigador ha de controlar o igualar la influencia de las restantes fuentes de variación que puedan afectar a la variable dependiente. Procurará que influyan por igual en el grupo control y en el experimental. Si todas las variables extrañas influyen por igual en ambos grupos, las diferencias (varianza) encontradas entre el grupo control y el experimental serán debidas a la variación introducida deliberadamente (niveles de VI).

Supongamos que queremos comprobar la eficacia de un programa basado en la utilización de medios audiovisuales (MA V) para la enseñanza del inglés en primaria. Para ello, seguimos las fases de la aleatorización, y después de comprobar la equivalencia inicial de los grupos (pretest), aplicamos el programa con MA V a un grupo experimental, y a un grupo control le aplicamos el programa sin MA V. Posteriormente, con una prueba paralela al pretest, medimos el rendimiento final en lengua inglesa (postest) y obtenemos las medias especificadas en la tabla 5.1.

TABLA 5.1 Ejemplo de diseño experimental para estudiar la eficacia de los MA V

Grupo	*n*	*Asignación*	*Pretest*	*Programa*	*Postest*
Experimental (GE)	5	Azar	2	Con MAV	9
Control (GC)	5	Azar	2	Sin MAV	5

Con las medias obtenidas podemos efectuar distintas comparaciones. Imaginemos que después de aplicar las pruebas estadísticas adecuadas obtenemos los resultados que aparecen en la tabla 5.2.

TABLA 5.2 Algunas comparaciones que pueden efectuarse en un diseño experimental

Objetivo	*Prueba estadística*	*Medias*	*Decisión*
Equivalencia inicial	t datos independ.	2 - 2	Ha
Equivalencia final (eficacia MA V)	t datos independ.	9 - 5	Hj
Cambio en GE	t datos apareados	9 - 2	H,
Cambio en GC	t datos apareados	2 - 5	H,

Cuando el investigador pretende contrastar una hipótesis a través de un experimento puede proceder del siguiente modo:

1) Mantiene constantes o iguala las variaciones (varianza) que puedan producir todas las posibles influencias que procedan de las variables extrañas. Esto es, controlar la influencia de las variables extrañas.
2) Introduce una variación deliberada controlando (manipulando) la variable independiente.
3) Analiza la intensidad de la variación que se produce en VD como consecuencia de la variación introducida al aplicar distintos niveles de la variable independiente.

De ahí que las conclusiones de un experimento se basen en las variaciones, cambios o varianzas observadas en la variable dependiente a raíz de las variaciones o cambios introducidos deliberadamente en la variable independiente, procurando igualar o sistematizar la varianza debida a variables extrañas. Para entender mejor la importancia de la varianza la analizaremos más detenidamente en el contexto del experimento planteado sobre los MA V.

En la tabla 5.3 se han esquematizado, de forma ficticia y aproximada, algunas fuentes de variación que podrían intervenir en el experimento de los MA V. Las puntuaciones en inglés obtenidas por los sujetos en el pretest presentan ciertas diferencias o variaciones que pueden atribuirse a tres fuentes fundamentales: diferencias individuales (motivación, inteligencia...), influencias ambientales (profesores, familia...) e imperfecciones de los instrumentos de medida. A pesar de esta varianza observada en las puntuaciones de los sujetos los grupos que se han formado tienen la misma media en inglés. Es decir, puede existir cierto grado de variación (debería constatarse estadísticamente) entre los sujetos que forman cada grupo (varianza intragrupo) sin que aparezcan necesariamente diferencias entre los grupos (varianza intergrupos).

Al iniciar el experimento, el investigador introduce una variación: utiliza los MA V en el grupo experimental, pero no en el grupo control. Ahora bien, como puede apreciarse en la tabla 5.3, otras variables como el profesor y la memoria de los sujetos también pueden originar variaciones.

El posible efecto de todas las fuentes de variación mencionadas queda recogido en la varianza total que presentan las puntuaciones en el postest de todos los sujetos. Como puede apreciarse en la tabla 5.3, la variación total de las puntuaciones de inglés se debe a unas variaciones constantes o sistemáticas que constituyen la denominada *varianza sistemática* y a otras irregulares o aleatorias que constituyen la *varianza residual o del error.*

A) *Varianza sistemática.* Es una variación originada por variables que desvían los datos en una dirección más que en otra, aumentando o disminuyendo, pero siempre en la misma magnitud y dirección. Esta desviación, en el experimento perfectamente controlado, es debida exclusivamente a la variable independiente (MA V). Cuando el grado de control no es el ideal, la desviación ha podido ser provocada también por variables extrañas que no han sido convenientemente controladas. Así, en el ejemplo propuesto el entusiasmo del profesor ha influido más en el grupo experimental (+ 2) que en el de control (+ 1).

En cambio, el hecho de que a los sujetos de ambos grupos les hayan proyectado las mismas películas en inglés durante las clases, tiene una incidencia idéntica tanto en el grupo control como en el experimental, es decir, esta influencia extraña ha sido controlada y deja de ser una fuente de variación. Al influir por igual (+ 1) no añade ninguna diferencia o varianza, no afecta a la variación existente entre los datos que resultan de medir VD. Por tanto, si queremos ver la influencia de los MA V debemos controlar las otras fuentes de variación.

TABLA 5.3 Posibles fuentes de variación que pueden intervenir en el experimento de los MA V (Los datos son ficticios; sólo se exponen con fines didácticos para facilitar la comprensión de los conceptos; por tanto, se omite la significación estadística)

	(VD) Pretest	*(VI) MAV*	*Profesor*	*Memoria*	*Error medida*	*Películas*	*(VD) Postest*
GC	2	(+0)	(−1)	(+4)	(−1)	(+1)	7
	1	(+0)	(+1)	(+2)	(−2)	(+1)	3
	3	(+0)	(+1)	(+2)	(−2)	(+1)	5
	2	(+0)	(+1)	(−1)	(−3)	(+1)	6
	2	(−0)	(+1)	(−2)	(−2)	(+1)	4
	$\bar{X} = 2$						$\bar{X} = 5$
GE	1	(+3)	(+2)	(−1)	(+1)	(+1)	7
	4	(+3)	(+2)	(+2)	(−1)	(+1)	11
	1	(+3)	(+2)	(+1)	(+2)	(+1)	10
	2	(+3)	(+2)	(+1)	(−1)	(+1)	8
	2	(+3)	(+2)	(−1)	(+2)	(+1)	9
	$\bar{X} = 2$	Varianza primaria	Varianza secundaria	Varianza intragrupo o del error			$\bar{X} = 9$

↑ ↑ ↑ ↑ ↑ desviación sistemática de los datos ↓ ↓ ↓ ↓ ↓

Varianza sistemática | Varianza residual

El propósito principal del investigador es separar aquellas fuentes de variación en las que se halla interesado (VI) de las que no le interesan (variables extrañas). De ahí la necesidad de considerar dos fuentes de variación sistemática: la varianza primaria y la varianza secundaria.

- *Varianza primaria.* Desviación sistemática de los datos debida a la manipulación de VI. Cuanto más efectiva sea VI, más diferencia o desviación aparecerá entre la medición de VD efectuada en el grupo control y en el experimental. De ahí que se denomine *varianza intergrupos.*
- *Varianza secundaria.* Desviación sistemática de los datos debida a otras variables que no son VI y no han sido controladas. Por ello influyen de forma desigual, provocando varianza no pretendida.

B) *Varianza del error o aleatoria.* Constituida por el conjunto de fluctuaciones que presentan las medidas de la variable dependiente que tienden a actuar al azar. Posibles fuentes son:

- *Diferencias individuales.* Constituidas por características inherentes a los sujetos; por ejemplo, en el experimento de los MA V podría ser la memoria de cada sujeto. Ante un mismo tratamiento experimental no todos los sujetos reaccionan igual. Dado que las diferencias individuales tienden a desviar los datos en sentidos dispares, esta varianza se considera debida al error o aleatoria por ser impredecible.
- *Errores de medida.* Puede deberse a instrumentos poco precisos.
- *Procedimiento experimental.* La forma de aplicarlo puede ser diferente para los distintos sujetos.
- *Instrucciones.* Pueden darse de forma diferente por distintos experimentadores.

5.2 PRINCIPIO FUNDAMENTAL DE LA VARIANZA

En el experimento planteado sobre los MA V comprobaremos el siguiente principio fundamental de la varianza, es decir,

Varianza total = Varianza intergrupos + Varianza intragrupos

Como se recordará, los resultados medios de inglés, obtenidos en el postest, eran los siguientes:

Sin MAV			*Con MAV*
G.C.	Inglés	G	Inglés
a	7	f	7
b	3	9	**11**
e	5	h	10
d	6	i	8
e	4	j	9
	$\bar{X}=5$		$\bar{X}=9$

Sucesivamente procederemos a realizar el cómputo de la varianza total y de las varianzas intergrupo e intragrupo, con el [m de apreciar que aquélla se puede desglosar en la suma de estas últimas.

Cálculo de la varianza total. Variación de cada puntuación con respecto a la media total. Es decir:

$$S^2 = \frac{\Sigma (X - \bar{X})^2}{n} = \frac{60}{10} = \boxed{6}$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{70}{10} = 7$$

	X	$(X - \bar{X})$	$(X - \bar{X})^2$
	7	0	0
Grupo control	3	−4	16
	5	−2	4
	6	−1	1
	4	−3	9
	7	0	0
Grupo experimental	11	4	16
	10	3	9
	8	1	1
	9	2	4
Σ	70		60

Cálculo de la varianza intergrupos (primaria). La varianza total se ha calculado a partir de las diez puntuaciones en lengua inglesa; en cambio, la varianza intergrupo se refiere a la variación entre las medias de los grupos. En consecuencia, se utiliza la misma fórmula, pero en el cálculo sólo intervienen las medias 5 y 9 representando a cada grupo:

$$S^2 = \frac{\Sigma (X - \bar{X})^2}{n} = \frac{8}{2} = \boxed{4}$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{14}{2} = 7$$

	X	$(X - \bar{X})$	$(X - \bar{X})^2$
Grupo control	5	−2	4
Grupo experi.	9	+2	4
Σ	14		8

Cálculo de la varianza intragrupo (del error). Promedio de las variaciones calculadas dentro de cada grupo con respecto a las medias respectivas. El cómputo sería el siguiente:

$$S_c^2 = \frac{\Sigma (X - \bar{X})^2}{n} = \frac{10}{5} = \boxed{2}$$

$$\bar{X} = \frac{\Sigma x}{n} = \frac{25}{5} = 5$$

	X	$(X - \bar{X})$	$(X - \bar{X})^2$
Grupo control	7	−2	4
	3	−2	4
	5	0	0
	6	+1	1
	4	−1	1
Σ	25		10

De esta manera queda comprobado que se cumple el principio fundamental de la varianza:

$$S_e^2 = \frac{\Sigma (X-\bar{X})^2}{n} = \frac{10}{5} = 2$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{45}{5} = 9$$

	X	$(X-\bar{X})$	$(X-\bar{X})^2$
Grupo experimental	7	−2	4
	11	+2	4
	10	+1	1
	8	−1	1
	9	0	0
Σ	45		10

$$S^2 = \frac{S_c^2 + S_e^2}{n} = \frac{2+2}{2} = \boxed{2}$$

De esta manera queda comprobado que se cumple el principio fundamental de la varianza:

6	=	4	+	2
Varianza total	=	Varianza intergrupos	+	Varianza intragrupos
Varianza total	=	Varianza sistemática	+	Varianza del error
Varianza total	=	Varianza primaria (VI) + Varianza secundaria (Extrañas)	+	Varianza residual (Dif. individuales errores)

Las pruebas de significación, como el análisis de la varianza, se basan en el cociente:

$$F = \frac{\text{Varianza intergrupos}}{\text{Varianza intragrupos}} = \frac{4}{2} = 2$$

En la medida que aumente la varianza intergrupos (refleja el efecto de VI) y disminuya la varianza intragrupos (diferencias individuales y errores) el experimento será más eficaz y sensible, pues el cociente aumentará. Si nuestro experimento ficticio estuviera perfectamente controlado se verificaría que:

varianza secundaria = 0

pues las variables extrañas influirán por igual en cada grupo o se habrá eliminado su influencia. Por tanto, en el experimento ideal,

varianza intergrupos = varianza primaria = 4 (Sistemática) (VI)

Por tanto, en el buen experimento:

varianza intergrupo > varianza intragrupo

Para ello es necesario un control que consistirá en:

1) Maximizar la varianza primaria (producida por VI).
2) Neutralizar o eliminar la varianza secundaria (producida por las variables extrañas).
3) Minimizar la varianza del error (producida por las diferencias individuales y los errores de medida).

5.3 CONTROL DE LA V ARIANZA

El experimento es un tipo especial de investigación para examinar relaciones de causalidad. Para ello se introducen cambios deliberados en la variable independiente y se registran posibles modificaciones en la variable dependiente.

Con el fin de garantizar la validez del experimento, se utilizan técnicas de control de la varianza que pueden consistir en manipulación física, selectiva o estadística:

Manipulación física. Son técnicas de control directo que pueden ejercerse al aplicar la variable independiente o al igualar los efectos de algunas variables extrañas. Por ejemplo, espejos unidireccionales, suministrar drogas, premios, presentar a los sujetos estímulos verbales, luminosidad, ruidos o temperatura.

Manipulación selectiva. Algunos ejemplos pueden ser seleccionar y asignar aleatoriamente los sujetos, mantener constantes ciertas condiciones, aparear los sujetos, el bloqueo, el balanceo y el contrabalanceo.

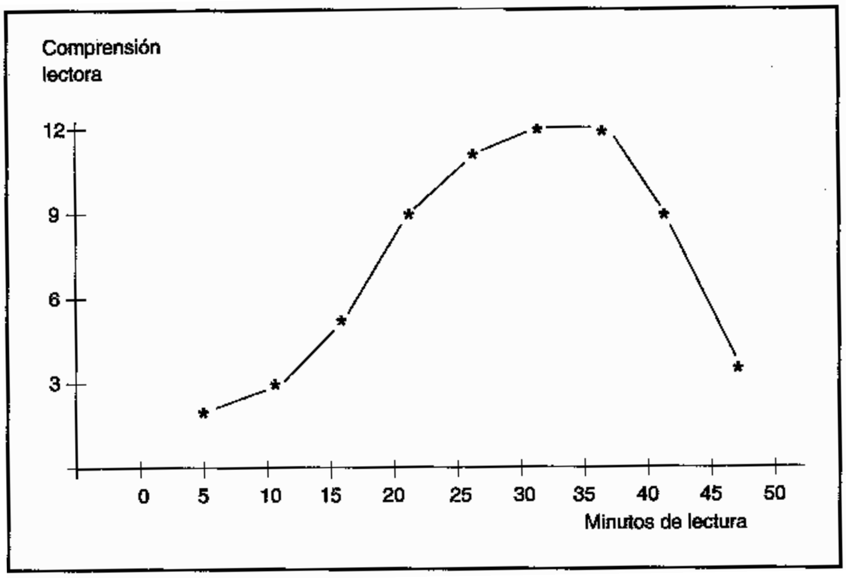
Manipulación estadística. Dada la naturaleza del fenómeno educativo es frecuente tener que recurrir a técnicas indirectas de control como son las técnicas estadísticas. Estas técnicas de control, como la correlación parcial y el análisis de covarianza, se utilizan cuando las variables no pueden ser sometidas a manipulación física o selectiva.

Como ya hemos apuntado más arriba, el investigador ha de recurrir a estas técnicas de control con el fin de maximizar la varianza primaria, eliminar la varianza secundaria

y minimizar la varianza del error. A continuación veremos más detenidamente distintas técnicas de control:

MAXIMIZACIÓN DE LA V ARIANZA PRIMARIA

Los niveles de la variable independiente han de adoptar en la realidad los valores decididos por el experimentador. Además, dichos valores han de ser extremos y óptimos. Supongamos que se realizan distintos experimentos en tomo a la eficacia de técnicas de estudio en niños de nueve años. Si queremos ver la influencia del tiempo de lectura en el rendimiento, habrá que asignar los valores adecuados. Así, valores como O, 1 Y 2 minutos de lectura pueden llevamos a concluir que el tiempo de lectura no influye, ya que no son valores extremos ni óptimos. Lo mismo podría ocurrir con valores más extremos, pero no óptimos, como O, 50 y 90 minutos, ya que al visualizar la relación entre dicha variable y una característica decisiva para el rendimiento (comprensión) se obtiene la figura 5.1.



FIG, 5.1. Distribución de la comprensión lectora según los minutos de lectura

En cambio, valores como O, 25 y 50 podrían ejercer un efecto más diferenciado, caso de existir una relación real. En consecuencia, podrían ser valores extremos y óptimos.

Existen tres formas de asignar los niveles de la variable independiente a los sujetos que permiten distintas modalidades de control: intergrupo, intragrupo y mixta.

Asignación intergrupo

Se utiliza esta modalidad cuando cada nivel de la variable independiente, también denominado *condición experimental* o *tratamiento,* se aplica a un grupo diferente de sujetos, es decir, cada sujeto es sometido a una sola condición experimental. Así, por ejemplo, para ver cómo influye el tiempo de estudio en la comprensión lectora de un texto T podemos seleccionar al azar 30 alumnos de B UP y asignarlos, también al azar, a un grupo control (GC) y dos grupos experimentales (GEl y GEl). Los alumnos asignados al GC no dedican tiempo a estudiar el texto, mientras que en cada GE se dedica 1 y 2 horas. Después de realizar el proceso descrito se comparan las medias de comprensión lectora referidas a dicho texto (tabla 5.4).

TABLA 5.4 Ejemplo de asignación de sujetos según la modalidad intergrupo

Grupo control (O horas)	*Grupo experimental* 1 (1 hora)	*Grupo experimental* 2 (2 horas)
1.Juan	11. Rosa	21. Lourdes
2. Ana	12. José	22. Carlos
3. Antonio	13. Raquel	23. Miguel
……….....	...	...
9. Ramón	19. Mireia	29. Sergio
10. María	20. Víctor	30. Mónica
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$

En este experimento, comparar las medias tiene sentido si los tres grupos que intervienen y las tres situaciones ya eran homogéneas o equivalentes, antes de empezar el experimento y durante el desarrollo del mismo, especialmente en lo que se refiere a diferencias que pueden ser ocasionadas por variables extrañas que pueden influir en la variable dependiente. Para conseguir que estas variables afecten por igual a los tres grupos de sujetos y hacerlos comparables hay que controlar o igualar en cada grupo influencias como las siguientes:

- El experimentador: entusiasmo, amabilidad, etc.
- El contexto experimental: ruido, local, hora, instrucciones, etc.
- Los sujetos experimentales que presentan diferencias individuales en motivación, inteligencia, edad, sexo, capacidad de estudio, etc.

Se consigue un control ideal cuando lo único que diferencia a cada grupo es el tratamiento proporcionado a través de los niveles de la variable independiente (O horas, 1 hora y 2 horas). En consecuencia, si antes de empezar el experimento los tres grupos presentaban medias estadísticamente iguales en la comprensión lectora del texto T,

puede concluirse, con más seguridad, que si aparece alguna diferencia entre las medias después del experimento, dicha diferencia podrá ser atribuida a que los grupos han sido sometidos a distintos niveles de la variable independiente.

Asignación intragrupo

Según esta modalidad todos los niveles de la variable independiente se aplican sucesivamente al mismo grupo de sujetos, es decir, cada sujeto es sometido a todas las condiciones experimentales. En el ejemplo anterior, podemos someter a los 30 sujetos a la prueba de comprensión lectora sobre el texto T para obtener la media cuando el nivel de la variable independiente vale O horas de estudio. En dos días sucesivos el grupo total estudia durante 1 hora y 2 horas respectivamente, midiéndose la comprensión después de cada período de estudio (tabla 5.5).

TABLA 5.5 Ejemplo de asignación de sujetos según la modalidad intragrupo

Grupo total (0 horas)	*Grupo total* (1 hora)	*Grupo total* (2 horas)
1. Juan	1. Juan	1. Juan
2. Ana	2. Ana	2. Ana
3. Antonio	3. Antonio	3. Antonio
...	...	...
...		...
29. Sergio	29. Sergio	29. Sergio
30. Mónica	30. Mónica	30. Mónica
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}$

En este caso no hay tanta necesidad de controlar variables extrañas que proceden de los sujetos. Al ser los mismos, dichas variables afectarán por igual a los resultados en comprensión lectora. En cambio, hay que controlar las variables vinculadas al experimentador y , especialmente, las que están implicadas en el contexto experimental y son inherentes a esta modalidad de asignación de los tratamientos. Así, hay que tener en cuenta que:

1) Los efectos de los niveles de la variable independiente (tiempo de estudio) pueden permanecer desde una aplicación hasta las siguientes acumulándose. En este sentido, para controlar este efecto no pretendido podría recurrirse a tres textos de dificultad similar pero con diferente contenido.

2) Que los sujetos pueden mejorar su comprensión lectora por el hecho de pasarles tres pruebas sobre comprensión.

3) Los niveles de VI pueden tener diferente efecto según el orden de aplicación.

Aunque existen técnicas de control que tratan de igualar la influencia de estas variables extrañas, para el experimento planteado es más viable una aplicación intergrupo; sin embargo, el mismo ejemplo ha sido incluido aquí con fines meramente expositivos.

Asignación mixta

En esta modalidad a varios sujetos se les aplica solamente un nivel de la variable independiente (intergrupos) y a otros sujetos se les aplican varios niveles de la variable independiente (intragrupo). Así, en el ejemplo mencionado, después de aplicar la prueba de comprensión al grupo total, el mismo grupo de 30 sujetos puede estudiar durante 1 hora, mientras que sólo 10 sujetos estudian durante 2 horas (tabla 5.6).

TABLA 5.6. Ejemplo de asignación de sujetos según la modalidad mixta

Grupo total (O horas)	*Grupo total* (1 hora)	*Grupo total* (2 horas)
1. Juan	1. Juan	1. Rosa
2. Ana	2. Ana	2. José
3. Antonio	3. Antonio	3. Raquel
...	...	...
...		
29. Sergio	29. Sergio	29. Mireia
30. Mónica	30. Mónica	30. Víctor
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$

Esta manera de asignar los tratamientos combina las ventajas de las condiciones intergrupo e intragrupo.

ELIMINACIÓN DE LA V ARIANZA SECUNDARIA

El experimentador ha de garantizar que los cambios observados en la variable dependiente (varianza) sean únicamente debidos a los distintos niveles de la variable independiente. Para ello las variables extrañas no han de contaminar los resultados finales del experimento, es decir, las variables extrañas han de incidir con la misma intensidad en los grupos experimentales y en el grupo control. De esta forma, los grupos experimentales y el grupo control sólo se diferenciarían en que reciben distintos niveles de la variable independiente. En consecuencia, las diferencias observadas en la variable dependiente serán atribuibles a los niveles de la variable independiente siempre que las variables extrañas no influyan de forma diferenciada.

Si el experimento no se controla adecuadamente, existe la posibilidad de que los

cambios observados en la variable dependiente (varianza total) puedan ser provocados, en parte, por una gran multiplicidad de variables extrañas que junto a la variable independiente pueden producir modificaciones en la variable dependiente. Las variables extrañas que pueden afectar a los resultados finales han sido englobadas en categorías como las siguientes:

- Sujetos experimentales: edad, sexo, capacidad cognitiva, personalidad, destreza, etc.
- Contexto experimental: instrucciones, aparatos, ruidos, etc.
- Experimentador: expectativas, actitudes, *rol,* etc.

De ahí que el investigador deba cerciorarse de que las variaciones que aparecen en la variable dependiente puedan ser atribuidas exclusivamente a las variaciones introducidas deliberadamente a través de la variable independiente. Para ello puede utilizar distintas técnicas de control.

Las principales formas y técnicas de *control de las variables extrañas* se basan fundamentalmente en la *aleatorización* y la selección. La *aleatorización* constituye el procedimiento más adecuado para controlar las variables extrañas conocidas y desconocidas que proceden de los sujetos. Gracias a la *selección* un mismo método didáctico puede utilizarse para explicar a los alumnos diferentes conceptos que pueden seleccionarse según su dificultad o adecuación a la edad. Es posible formar grupos según el sexo, edad, clase social, motivación, inteligencia o personalidad cuando sospechamos o sabemos que estas variables pueden tener una influencia significativa. Mediante la selección se pueden eliminar variables, como pueden ser el ruido, o bien elegir un valor único o una única categoría: niños o niñas, período evolutivo, procedencia, etc., teniendo en cuenta que limitaremos el poder generalizador de las conclusiones. Algunas variables contextuales, como la temperatura y la hora de aplicación de los instrumentos de medida, también podemos mantenerlas constantes.

A continuación se detalla la técnica de la aleatorización y diversos procedimientos, algunos de ellos basados en la selección:

Aleatorización

Consiste en asignar los sujetos al azar a cada grupo del experimento, lo que permite igualar todas las diferencias individuales (carácter, inteligencia, motivación, comprensión lectora, etc.). La asignación aleatoria asegura que dichas variables tengan la misma oportunidad de influir en la variable dependiente, llegando a compensarse sus efectos y resultando grupos equivalentes antes de aplicarse la variable independiente.

La aplicación estricta de esta técnica requiere las siguientes fases (fig. 5.2):

1) Seleccionar al azar una muestra de la población.

2) Asignar los sujetos al azar a tantos grupos como niveles tenga la variable independiente. Por ejemplo, grupo 1 (G_1) y grupo 2 (G_2).

3) Formados los grupos del experimento se asigna al azar cada nivel de la variable independiente a un grupo. El grupo que reciba el nivel mínimo de la variable independiente o ausencia de tratamiento experimental (T_o), se denominará grupo control

(GC), Y grupo experimental (GE) el que reciba el tratamiento (TI). De ahí que estas dos últimas fases sean más propias de la modalidad de asignación intersujetos.

Cuando el número de sujetos de la muestra es suficientemente grande, la probabilidad de que las diferencias individuales se compensen unas a otras, por efecto del azar, es también mayor, existiendo más garantía de que los grupos sean equivalentes antes de iniciarse el experimento. Como es evidente, aquellas variables extrañas que puedan influir o generarse después de haberse formado los grupos (cansancio, ansiedad, experimentador, ambiente, contexto experimental, etc.) no pueden controlarse mediante la técnica de la aleatorización.

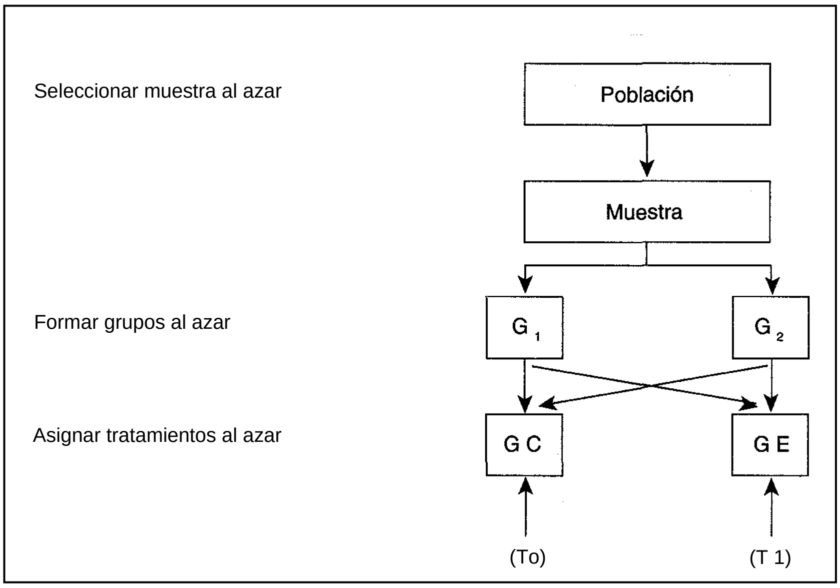


FIG. 5. 2 Fases de la aleatorización como técnica de control (adaptado de Rosel, 1986,244)

Constancia

Las variables extrañas se mantienen a un nivel constante. Por ejemplo, si queremos ver la eficacia de distintos métodos didácticos se procurará que para todos los grupos del experimento los temas que hay que explicar sean los mismos. También convendrá que otras características, como el nivel de iluminación del local, la claridad al especificar las instrucciones, etc., intervengan por igual en cada grupo del experimento. La constancia (y también la eliminación) presenta el inconveniente de que restringe la generalización de las conclusiones del experimento. Así, cuando la edad se mantiene constante a los 13 años o cuando sólo se seleccionan chicos, los resultados del experimento no serán extensibles a los alumnos de 10 años ni a las chicas, respectivamente.

Eliminación

Es un caso especial de constancia que se lleva a cabo cuando algunas condiciones ambientales se restringen al máximo, manteniéndolas constantes pero a nivel de ausencia, es decir, a nivel cero. Por ejemplo, pueden eliminarse elementos distorsionantes como el ruido, la visión de objetos que distraen, etc., lo que hace que ignoremos qué ocurriría si el experimento se llevara a cabo en condiciones normales.

Apareo

Cuando sospechamos que una variable extraña (variable de apareo) está correlacionada y/o puede influir decisivamente en la variable dependiente, podemos formar grupos equivalentes antes de aplicar la variable independiente, recurriendo al apareo de los sujetos. Para ello asignamos a cada grupo del experimento un sujeto que posea la misma magnitud en la variable de apareo o de bloqueo.

Supongamos que queremos comprobar la eficacia de un programa de informática para resolver ecuaciones. Podemos utilizar la inteligencia como variable de apareo, después de comprobar que se halla fuertemente correlacionada con la habilidad para resolver ecuaciones. En la práctica, aunque se utilice la aleatorización, cuando la muestra no es muy grande, puede darse el caso de que el grupo experimental esté constituido por sujetos más inteligentes. Si así ocurriera, el posible aumento en la habilidad para resolver las ecuaciones podría deberse no sólo al programa de informática, sino también a la inteligencia. Así, para aplicar la técnica del apareo en el ejemplo propuesto podríamos aplicar el test de inteligencia «D-48», lo que nos permitiría formar un grupo control (GC) y otro experimental (GE), agrupando por parejas los sujetos más afines en dicha variable y asignándolos al azar a cada grupo, lo que origina distintos pares o bloques (tablas 5.7 y 5.8).

Como los sujetos dentro de cada bloque o par tienen la misma inteligencia, resultará que los grupos control y experimental serán equivalentes antes de aplicar el programa de informática. Como es evidente, la propia variable dependiente también podría utilizarse como variable de apareo.

TABLA 5. 7 Puntuaciones de inteligencia «D-48»

1. Eva	30	7. Ignacio	35
2. Pablo	33	8. Rosa	28
3. Carlos	34	9. Javier	33
4. Gemma	35	10. Ana	30
5. Manuel	28	11. Marco	27
6. Carmen	27	12. Sonia	34

TABLA 5.8 Pares de sujetos y asignación a cada grupo

Pares	GC	*GE*
1.	Ignacio	Gemma
2.	Carlos	Sonia
3.	Pablo	Javier
4.	Eva	Ana
5.	Rosa	Manuel
6.	Carmen	Marco
	X_0	X_j

En la representación simbólica de un diseño basado en la técnica del apareo los distintos pares pueden representarse por:

Ap = asignación según apareo

En general, esta técnica de control presenta inconvenientes como los siguientes:

- Cuando la variable de apareamiento presenta gran variabilidad hay que tener más flexibilidad en el grado de semejanza que han de tener las puntuaciones para incluidas en el mismo par.
- Con frecuencia, hay que prescindir de algunos sujetos si no tienen un «par» adecuado que presente un nivel similar en la variable de apareamiento.
- De la misma manera, cuando ya se han formado los pares y falla un sujeto por enfermedad o absentismo, hay que eliminar el par si no existe otro sujeto similar que pueda sustituir al ausente.
- Si se quiere aparear a los sujetos a partir de más de una variable de apareo, las posibilidades de formar los pares son menores y la muestra quedaría muy reducida.

Bloqueo

Cuando se forman más de dos grupos de tratamiento, en lugar de agrupar por pares interesa formar bloques a partir de una o más variables de bloqueo. Esta técnica presenta inconvenientes similares a los que se han expuesto al considerar la técnica del apareo y hay que admitir un mayor margen de variabilidad en los sujetos que se agruparán en un mismo bloque; de lo contrario habría que prescindir de muchos sujetos. Sin embargo, como dentro de cada bloque los sujetos se asignan también al azar a cada grupo de tratamiento, es fácil conseguir grupos equivalentes con respecto a la variable de bloqueo.

En el ejemplo planteado en el apartado anterior, si queremos comprobar la eficacia relativa de tres programas de informática para resolver las ecuaciones, podemos formar un grupo control y tres grupos experimentales, utilizando las puntuaciones de inteligencia para agrupar tres bloques de inteligencia alta, media y baja (tabla 5.9).

Balanceo

Esta técnica distribuye equitativamente la presencia de una variable extraña en los grupos que deseamos formar. Si en el experimento mencionado anteriormente sobre la eficacia de los programas de informática se sospecha que las chicas pueden tener una mayor habilidad para la resolución de ecuaciones, no sabríamos si la posible mejora observada en los grupos experimentales es debida a los programas, al hecho de que en dichos grupos pudiera predominar el sexo femenino o a ambas circunstancias.

TABLA 5.9 Ejemplo ilustrativo de la técnica de bloqueo

Bloques autorizados	*GC*	*GE₁*	*GE₂*	*GE₃*
1. Inteligencia alta 2. Inteligencia media 3. Inteligencia baja	Ignacio Eva Carmen	Carlos Javier Marco	Gemma Ana Rosa	Sonia Pablo Manuel
	$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$

Supongamos que al comparar las medias se obtiene que bajo el programa 2 (GE2) la habilidad para resolver las ecuaciones es superior a la que presentan los sujetos que han utilizado el programa 1 (GE1). Esta diferencia entre las medias obtenidas podría deberse a que el programa 2 es más eficaz que el programa 1 y/o al hecho de que el grupo GE2 está constituido por chicas y el grupo GE1 por chicos. En consecuencia, los resultados del experimento quedarían contaminados por la variable sexo. Hay que tener en cuenta que la muestra debería ser más grande y entonces sería menos probable que predominara uno de los sexos en algún grupo después de aleatorizar la asignación dentro de cada bloque. Una posible solución sería formar los grupos incluyendo la misma cantidad de chicos y chicas en cada grupo. Para recoger la mayor información posible con respecto a los efectos de la variable sexo, y suponiendo que no interese controlar la inteligencia, el experimento podría diseñarse según la técnica del balanceo para controlar la variable sexo. Para una mayor sencillez veamos cómo podría comprobarse la eficacia de un solo programa (tabla 5.10).

TABLA 5. 10 Ejemplo con la utilización de la técnica del balanceo

		GC	GE	
Sexo	M	Ignacio Carlos Pablo	Javier Marco Manuel	$\bar{X}_M$
	F	Eva Carmen Gemma	Ana Rosa Sonia	$\bar{X}_F$
		$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_T$

Como puede observarse, el experimento tiene cuatro subgrupos de sujetos: dos subgrupos forman el grupo control (tres chicos y tres chicas) y el programa es aplicado a otros dos subgrupos (tres chicos y tres chisas) y dos subgrupos. Si se incluyen más sujetos en cada subgrupo, la comparación $\bar{X}_0$ y $\bar{X}_1$ nos informa sobre el efecto del programa, ya que en el grupo control no ha sido aplicado. Además, podemos obtener una información adicional sobre el sexo al comparar las medias $\bar{X}_M$ y $\bar{X}_F$. En realidad, hemos sistematizado o equilibrado una variable extraña (sexo), considerándola como otra variable independiente de la investigación.

Contrabalanceo o equiponderación

Con esta técnica, los grupos se dividen en pequeños sub grupos con el fin de equilibrar los efectos de las variables extrañas. Puede utilizarse en la modalidad de asignación intragrupo, es decir, cuando distintos tratamientos aplicados sobre los mismos sujetos pueden producir efectos distintos según la secuencia u orden de aplicación. Supongamos que hemos elaborado varias reglas nemotécnicas para potenciar las facultades y alcance de la memoria de un grupo de niños débiles mentales y pretendemos comprobar la eficacia relativa de dos técnicas (T) y (T) para memorizar frases. Como es obvio, al aplicar cada técnica (VI) y al contabilizar cuántas frases han memorizado (VD), se recurriría a frases distintas pero de dificultad similar. Sin embargo, en cada aplicación habría que tener en cuenta dos variables extrañas que pueden generarse durante el experimento:

1) Cierto nivel de aprendizaje práctico, experiencia o habilidad que puede facilitar la memorización de las frases después de aplicar la primera técnica.
2) La fatiga, que puede favorecer que el sujeto cometa más errores a medida que aumenta el número de tratamientos.

El grado de memorización de las frases dependerá no sólo de la eficacia de la técnica aplicada, sino también de la acción conjunta del aprendizaje y de la fatiga, que originan el denominado *error progresivo.* Aunque éste puede disminuir con adecuados intervalos intertratamientos, ya que el tiempo suele eliminar los efectos de dicho error progresivo, es más viable neutralizarlo por medio del contrabalanceo (tabla 5.11).

TABLA 5. 11 Técnica de contrabalanceo. Modalidad intragrupo

T_1 (1.ª)	T_2 (2.ª)	T_2 (3.ª)	T_1 (4.ª)
Alicia	Alicia	Javier	Javier
Iván	Iván	Isabel	Isabel
Roberto	Roberto	Laura	Laura
...	...	...	...
...	...	...	...
Olga	Olga	Irene	Irene
Ramón	Ramón	Vicente	Vicente
$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$

La equiponderación o contrabalanceo puede adoptar también la modalidad intergrupo (tabla 5. 12).

TABLA 5. 12 Técnica de contrabalanceo. Modalidad intergrupo

T_1 (1.ª)	T_2 (2.ª)	T_2 (3.ª)	T_1 (4.ª)
Alicia	Alicia	Javier	Javier
Iván	Iván	Isabel	Isabel
Roberto	Roberto	Laura	Laura
...	...	...	...
...	...	...	...
Olga	Olga	Irene	Irene
Ramón	Ramón	Vicente	Vicente
$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$

En ambas modalidades se ha incluido entre paréntesis la secuencia de aplicaciones y la media del número de frases memorizadas después de haber practicado con cada técnica. Así, los niveles medios que permitirán estimar la eficacia relativa de cada regla nemotécnica podrían ser:

$$X_{T1} = \frac{\bar{X}_1 + \bar{X}_2}{2} \qquad X_{T2} = \frac{\bar{X}_3 + \bar{X}_4}{2}$$

ANÁLISIS DE COV ARIANZ

Las técnicas de control descritas pueden plantear al educador dificultades como las siguientes:

1) Formar nuevos grupos para el experimento, por ejemplo, a través de la aleatorización, puede ser no recomendable por razones organizativas de los centros. Esto obliga a utilizar grupos naturales, respetando las agrupaciones que constituyen las distintas clases o conglomerados.
2) La utilización de otras técnicas como el apareo y los bloques puede resultar inapropiada, ya que el educador suele disponer de pocos sujetos para formar grupos homogéneos igualados en variables extrañas relevantes.

En consecuencia, al tomar grupos ya constituidos (naturales), las variables extrañas pueden estar presentes con distintos niveles de intensidad, originando diferencias entre los grupos antes de empezar el experimento. Si estas diferencias iniciales no se pudieran controlar (igualar) mediante algún sistema alternativo de control, la investigación no podría realizarse. Por ejemplo, podría darse la situación ficticia de la tabla

5.13 al aplicar dos métodos de enseñanza (M_1 y M_2) a dos grupos de tercero de primaria y analizar sus efectos sobre el rendimiento tomando un grupo control (GC).

TABLA 5.13 Datos de una investigación en la que no se ha controlado la inteligencia

Grupos naturales	*Número sujetos*	*Inteligencia media (0-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
3.ºA	26	30	2	GC	4
3.º B	29	28	1	M_1	7
3.º C	27	32	3	M_2	8

Supongamos que la inteligencia está significativamente correlacionada con el rendimiento, y no podemos controlarla mediante las técnicas mencionadas. El análisis de covarianza permite un control estadístico o control indirecto de esta variable extraña al tener en cuenta las diferencias iniciales entre tercero A, tercero B y tercero C. Gracias al análisis de covarianza el experimento puede realizarse. Esta técnica estadística ajusta las diferencias finales con respecto a las iniciales y compara las medias tomando en consideración la posible existencia de diferencias iniciales que podrían enmascararlas. Es decir, permite corregir o «ajustar» los cambios observados en la variable dependiente, dejando solamente los efectos atribuibles a la variable independiente.

La situación planteada es comparable a lo que ocurriría con tres niños en un concurso para coleccionar cromos. Si el educador quiere saber quién es más hábil coleccionando y sospecha que los niños empiezan el concurso con distinto número de cromos, habrá de comparar las cantidades finales acumuladas tomando en consideración la cantidad inicial que ya tenía cada niño. Es decir, ha de «ajustar» o hacer comparables las cantidades finales. La pregunta implícita sería: ¿cuántos cromos habría ganado cada niño si hubieran empezado sin ninguno o con la misma cantidad?

En el caso que nos ocupa el análisis de covarianza resuelve estadísticamente este interrogante y compara las medias que tendría después cada grupo si inicialmente las hubieran tenido iguales. Por ejemplo, las medias definitivas comparables o ajustadas hubieran podido ser 4, 8 y 7, lo que podría sugerir cierta supremacía del método M 1 (tabla 5.14):

TABLA 5. 14 Datos ficticios de una investigación en la que se ha controlado la inteligencia ajustando las medias a través de un análisis de covarianza

Grupos naturales	*Número sujetos*	*Inteligencia media (0-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
6.º A EGB	26	30	2	GC	4
6.º B EGB	29	28	1	M_1	8
6.º C EGB	27	32	3	M_2	7

En cierto modo, la aportación del análisis de covarianza sería similar a la información que ofrecería el siguiente diseño si en el caso planteado se hubiera podido controlar la inteligencia aleatorizando o seleccionando grupos de sujetos equivalentes (tabla 5.15).

TABLA 5.15 Datos de una investigación en la que se ha controlado la inteligencia directamente aleatorizando o seleccionando grupos equivalentes

Grupos naturales	*Número sujetos*	*Inteligencia medía (D-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
3.º A EGB	26	30	2	GC	4
3.º B EGB	29	30	2	M_1	8
3.º C EGB	27	30	2	M_2	7

La variable extraña (inteligencia) que se distribuye de forma desigual en los grupos del experimento recibe el nombre de *covariable.* El análisis de covarianza se basa en el supuesto de que el nivel en rendimiento obtenido por cada sujeto depende del valor que tenga en inteligencia y del método que se le aplique. En consecuencia, la media en rendimiento se reajusta en función de los métodos de enseñanza y en función de la covariable inteligencia.

MINIMIZACIÓN DE LA V ARIANZA INTRAGRUPO

Como se recordará, la varianza del error, residual o intragrupo se refería a variaciones debidas al azar. También habíamos contemplado algunas fuentes de esta varianza, como son las diferencias individuales y los errores de medida debidos a la imprecisión de los instrumentos utilizados en el proceso de medición.

Para minimizar la varianza del error conviene utilizar instrumentos fiables al medir la variable dependiente, reducir las diferencias individuales, aplicar bien los niveles de la variable independiente y comunicar las mismas instrucciones a los sujetos que intervienen en el experimento.

5.4 FUENTES DE INVALIDEZ

Hemos visto que en las investigaciones hay que tener en cuenta tres tipos de validez: la interna, que exige el control; la externa (generalización) y la conceptual (coherencia). En el caso de la investigación experimental estos tipos de validez revisten una importancia decisiva. Las variables o elementos que ponen en peligro la validez del experimento se denominan *fuentes de invalidez.* Por su relevancia nos centraremos en las fuentes de invalidez interna y externa.

FUENTES DE INVALIDEZ INTERNA

Cuando un experimento carece de validez interna, el investigador no tiene garantías de que las diferencias observadas (varianza total) en la variable dependiente puedan atribuirse exclusivamente a la aplicación intencional de los distintos niveles de la variable independiente. Existe la posibilidad de que dicha variación sea debida a la influencia de otras variables extrañas o fuentes de invalidez, es decir, hay motivos para sospechar que las variables extrañas no influyen por igual en todos los grupos del experimento, ya que no se aplican técnicas de control adecuadas. En consecuencia, cualquier elemento que afecte al grado de control del diseño de investigación constituye una fuente potencial de invalidez interna.

Las variables extrañas, que de no controlarse pueden constituir fuentes de invalidez interna, las agruparemos bajo las siguientes denominaciones genéricas:

a) *Historia.* Engloba cualquier acontecimiento externo a los sujetos y diferente de la variable independiente que influya a lo largo del experimento (entre el pretest y el postest) en la variable dependiente. Estos elementos distorsionantes pueden producirse dentro o fuera del marco del experimento. El efecto puede agudizarse cuando

1) el tratamiento es largo;
2) se trata de un experimento de campo;
3) los grupos no reciben los tratamientos simultáneamente.

b) *Maduración.* Consiste en un cambio que se produce en los sujetos debido a su desarrollo biológico o psicológico al transcurrir el tiempo. Puede estimarse a través del cambio observado en un grupo control.

c) *Selección diferencial de los sujetos.* Puede ocurrir en investigaciones cuasiexperimentales, cuando no se forman al azar los distintos grupos que intervienen en el diseño. Así, por ejemplo, cuando el investigador ha de utilizar grupos ya constituidos como aulas de un colegio. Dado que en cada grupo las variables extrañas han podido influir previamente de forma diferenciada, no puede garantizarse que los grupos sean homogéneos antes de aplicar la variable independiente. En consecuencia, los cambios observados pueden ser debidos a las diferencias existentes entre los grupos más que a la acción del tratamiento.

d) *Pérdida de sujetos.* Esta fuente de invalidez, también denominada *mortalidad experimental,* puede afectar especialmente si la pérdida se produce en los sujetos situados en los extremos de la distribución.

e) *Interacciones.* Pueden aparecer entre selección y maduración, selección e historia, etc. El grupo experimental puede hallarse más motivado, o ser más ágil al entender las instrucciones, y por ello, independientemente del tratamiento, puede obtener mayor puntuación que el grupo control en el postest.

f) *Reactividad de la medida.* Las medidas o la aplicación de pruebas puede afectar a los sujetos, o pueden responder éstos de forma diferente a como lo harían en condiciones normales, lo que influiría en los resultados finales, independientemente de la acción de la variable independiente.

g) *Instrumentación.* Ciertos instrumentos de medida pueden cambiar con el tiempo. Por ejemplo, algunos sujetos pueden hacer marcas en los protocolos de

algunas pruebas que condicionarán las respuestas de otros sujetos si se utilizan los mismos protocolos, o un observador puede cambiar de criterio durante la recogida de datos o durante el análisis de los mismos. También puede ocurrir que los instrumentos de medida no tengan una fiabilidad adecuada.

h) *Regresión estadística.* Cuando se aplica una prueba en sucesivas ocasiones (pretest-postest) las puntuaciones progresivas tienden a aproximarse a la media del grupo de procedencia. Este efecto es más pronunciado cuando se seleccionan grupos en base a puntuaciones extremas. Por ejemplo, si queremos probar la eficacia de un programa de técnicas de estudio para los mejores alumnos de un curso de primaria podemos aplicar un test de rendimiento con el fin de seleccionar los alumnos que presentan las puntuaciones más altas. Si se les pasa posteriormente un nuevo test de rendimiento, independientemente de la aplicación de las técnicas de estudio, la media que obtendrá el grupo en este segundo test tenderá a desplazarse hacia la media de la población original, es decir, tiende a descender. Si se hubieran seleccionado los alumnos más retrasados, la segunda vez que se pasara el test la media tendería a incrementarse. En ambos casos ocurre un fenómeno de regresión estadística, principalmente debido al azar y a posibles deficiencias de los instrumentos.

La regresión estadística produciría el efecto que aparece en la figura 5.3. Supongamos que al aplicar el primer test el grupo total obtiene una media global de 5, el grupo de alumnos más retrasado sólo registra una media de 3 y los más adelantados de 8. Si realizamos la segunda aplicación de un test de rendimiento, cubriendo el mismo contenido y sin introducir ningún cambio en los sujetos, la media del grupo más retrasado tenderá a aumentar, aproximándose a la media global del grupo de procedencia (5). Por el contrario, la media del grupo más adelantado tenderá a descender aproximándose a la misma media global. Así, si un grupo con puntuaciones altas en alguna variable se utiliza como grupo control, y otro con puntuaciones bajas se constituye en experimental, la regresión no está controlada y puede actuar favoreciendo a los sujetos experimentales.

i) *Difusión del tratamiento.* Los sujetos del grupo experimental pueden comunicar a los demás las características del tratamiento.

FUENTES DE INVALIDEZ EXTERNA

Las fuentes de invalidez externa son condiciones que afectan a la representatividad o generalización de los resultados de una investigación, es decir, a la posibilidad de generalizar los resultados a otros sujetos, a otros grupos y a otras condiciones y situaciones reales.

Las fuentes de invalidez externa más importantes que pueden afectar a la representatividad de los resultados son:

a) *Reactividad experimental.* Cuando los sujetos sometidos a las condiciones experimentales actúan de diferente forma de como lo hacen en condiciones normales. Esta variable puede afectar también a la validez interna.

b) *Interacciones.* Pueden producirse entre los errores de selección y la variable experimental. Cuando se realiza un experimento sobre una muestra sesgada (voluntarios, sujetos de un determinado nivel cultural, social, etc.), los resultados no podrán

generalizarse a otros grupos diferentes no representados en la muestra. De la misma manera, si el experimento se lleva a cabo en un contexto muy específico (laboratorio, colegio, centro de rehabilitación, etc.), no existen suficientes garantías de que los resultados puedan extrapolarse a otros ámbitos. En general, para que los resultados puedan generalizarse a situaciones de la vida real, el experimento ha de realizarse en condiciones similares a contextos reales.

c) *Efecto reactivo o interactivo del pretest.* La aplicación de un test previo puede restringir el alcance de la generalización de los resultados obtenidos. El pretest puede entrenar a los sujetos en relación con la variable independiente; así, los resultados obtenidos no serán generalizables a un grupo de sujetos que no hayan pasado el pretest.

d) *Inteiferendas de tratamientos múltiples.* Cuando se administra una serie de tratamientos, a menos que el efecto de los tratamientos anteriores quede eliminado con el tiempo, no se sabrá si los sujetos reflejan el efecto del tratamiento que se les aplica en ese momento o la posible interacción existente entre dicho tratamiento y los efectos de los tratamientos anteriores.

e) *Novedad del tratamiento.* La curiosidad, motivación o novedad pueden influir «aumentando» el efecto de la variable independiente. Cuando se aplicara el mismo tratamiento en un contexto real perdería efectividad a medida que se prolongara el tratamiento.

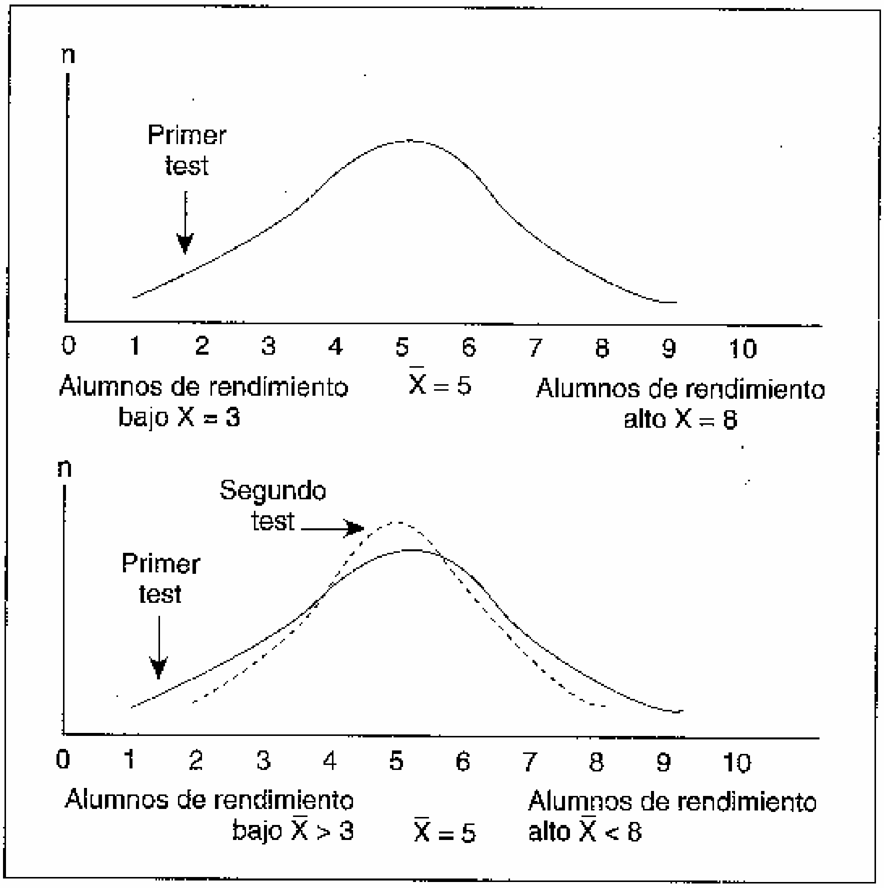


FIG. 5. 3 Efecto de la regresión estadística cuando se aplica un test por segunda vez en grupos extremos

5.5 CRITERIOS DE CLASIFICACIÓN DE LOS DISEÑOS

El método experimental puede adoptar diversas modalidades, siendo muchos los criterios de clasificación que pueden utilizarse: técnicas de control (Amau, 1978), número de grupos (Jiménez, 1985), mayor o menor rigurosidad, validez interna y externa (Campbell y Stanley, 1979), número de variables de tratamiento (Kirk, 1982). La adopción de uno u otro tipo de criterio genera taxonomías diferenciadas.

Didácticamente, el criterio del control parece ser el que inicialmente ofrece más ventajas, al desglosar los diseños en tres categorías: grupos al azar, bloques homogéneos e intragrupo. Dentro de cada categoría pueden efectuarse subdivisiones según el criterio del número de variables independientes y del número de niveles de tratamiento; así ocurre con los experimentos bivalentes, multivalentes y factoriales.

En la figura 5.4 pueden verse los diseños más importantes con asignación intergrupo o intragrupo y clasificados según el número de variables y las técnicas de control.

Para elegir el diseño que mejor se adapte a las necesidades del investigador y del problema planteado hay que tener en cuenta criterios como los siguientes:

1) Número de variables independientes que intervienen en la investigación. Cuando interviene una sola variable independiente hay que optar por los diseños unifactoriales o simples, y si están implicadas más de una variable independiente hay que elegir entre los multifactoriales o simplemente denominados factoriales.

2) Número de variables dependientes. Nuestra exposición se centrará siempre en diseños que contemplen una sola variable dependiente.

3) La asignación de los sujetos a los niveles de la variable independiente. Podemos elegir entre la asignación intergrupo e intragrupo. Si se opta por la asignación intergrupo podemos formar los grupos completamente al azar o según bloques. Como puede apreciarse en la figura 5.4, las diferencias individuales serán máximas en los diseños completamente al azar y mínimas en los diseños intragrupo.

4) Número de niveles de la variable independiente. Cuando dicha variable tiene dos niveles se utiliza un diseño bivalente, y si tiene más de dos se utiliza un diseño multivalente.

En los apartados siguientes se describen brevemente los diseños más utilizados.

5.6 DISEÑOS COMPLETAMENTE AL AZAR

Después de elegir la muestra al azar, en estos diseños los grupos se forman asignando al azar cada sujeto a uno de los grupos del experimento, y posteriormente, asignando también aleatoriamente cada nivel de tratamiento a los distintos grupos. Para que el azar pueda garantizar la equivalencia de los grupos hay que recurrir a muestras suficientemente grandes. De esta manera puede esperarse que las posibles variables contaminadoras se compensen y afecten por igual a todos los grupos antes de empezar los tratamientos.

Dentro de esta modalidad los diseños más utilizados son los bivalentes, multivalentes y factoriales.

Disminuyen las diferencias individuales →
Se reduce el error experimental

CRITERIOS / DISEÑOS	1.º Núm. VI	2.º Núm. VD	3.º ASIGNACIÓN EXPERIMENTAL: INTERGRUPOS – Completamente al azar	INTERGRUPOS – De bloques	INTRAGRUPOS
UNIFACTORIAL O SIMPLE	1	1	* Bivalente: Pretest - postest; Sólo postest * Multivalente: Pretest- postest; Sólo postest	* Grupos apareados * Bloqueo simple	* Tratamientos por sujetos * Análisis de tendecias
MULTIFACTORIAL O FACTORIAL	> 1	1	* Factorial	* Factorial bloques al azar	* Factorial intragrupo
* Análisis de la covarianza					

FIG. 5.4 Clasificación de los diseños más importantes. (Tomado de M. Bartolomé, 1980)

DISEÑOS BIV ALENTES

Se utilizan cuando aplicamos dos niveles de la variable independiente y generalmente se componen de un grupo experimental y un grupo control, o bien de dos grupos experimentales (sin grupo control). Cuando se utiliza el grupo control es posible equilibrar la influencia de las variables extrañas, pues tenderán a influir con la misma intensidad en ambos grupos. De esta manera muchas variables extrañas pueden quedar neutralizadas.

Supongamos que pretendemos detectar el efecto relativo del tiempo de estudio (A) sobre la asimilación de un contenido (VD). Si queremos analizar el efecto de 2 horas de estudio (a), y pasamos una prueba previa para comprobar la homogeneidad o equivalencia inicial y conocer el nivel de la variable dependiente antes de aplicar la independiente, utilizaremos un diseño pretest-postest (tabla 5.16) con grupo control. A este grupo le asignaremos el nivel mínimo de la variable independiente.

El grupo control es un grupo de referencia, pues, al igual que en el experimental, las mismas variables extrañas contextuales influirán con la misma intensidad en ambos grupos. En consecuencia, las diferencias observadas al medir la variable dependiente en cada grupo podrán atribuirse al tratamiento aplicado, siempre que los grupos sean inicialmente equivalentes. El grupo control permite estimar los cambios debidos a la

maduración $\bar{X}_1 - \bar{X}_2$ en aquellos experimentos suficientemente prolongados.

TABLA 5.16 Diseño bivalente pretest - postest

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	$\bar{X}_1$	a_0	$\bar{X}_2$
2 Experimental	25	Azar	$\bar{X}_3$	a_1	$\bar{X}_4$

El pretest puede aplicarse a la variable dependiente o a una variable extraña con la que puede estar relacionada. Si omitimos la primera prueba y asumimos la equivalencia inicial de los grupos, el diseño se denomina *sólo postest* (tabla 5.17). Con esta modalidad podemos comparar los niveles alcanzados por la variable dependiente en el grupo control y en el experimental, pero no podemos estimar los cambios operados en el grupo control y en el experimental pomo poder comparar los niveles antes y después del tratamiento.

TABLA 5.17 Diseño bivalente sólo postest

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	-	a_0	$\bar{X}_1$
2 Experimental	25	Azar	-	a_1	$\bar{X}_2$

DISEÑOS MULTIV ALENTES

Se utiliza cuando se aplican tres o más niveles de la variable independiente (tabla 5.18).

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	$\bar{X}_1$	a_0 (oh)	$\bar{X}_2$
2 Experimental	25	Azar	$\bar{X}_3$	a_1 (2h)	$\bar{X}_4$
3 Experimental	20	Azar	$\bar{X}_5$	a_2 (3h)	$\bar{X}_2$

Tanto en diseños bivalentes como multivalentes (pretest-postest), si se demuestra la equivalencia inicial (homogeneidad) de los grupos, bastará comparar las medias del

postest para comprobar el efecto relativo de las horas de estudio en el rendimiento (tabla 5.19).

TABLA 5.19 Posibles comparaciones que pueden efectuarse en diseños bivalentes y multivalentes

Objetivo	Comparación	Prueba
Homogeneidad o equivalencia inicial	$\bar{X}_1 - X_3 - X_5$	ANOVA
Homogeneidad final (prueba hipótesis, efecto VI)	$\bar{X}_2 \quad \bar{X}_4 \cdot \bar{X}_6$	ANOVA
Cambio en grupos experimentales	$\bar{X}_4 - \bar{X}_3$	t apareados
	$X_6 \cdot \bar{X}_5$	t apareados
Cambio en grupo de control (Influencia externa,	$\bar{X}_2 - \bar{X}_1$	t apareados

Si los grupos experimentales muestran un mayor cambio que el control puede inferirse que el tratamiento origina la diferencia.

Siempre que sea posible, es más adecuado utilizar diseños multivalentes, ya que los diseños bivalentes ofrecen una generalización muy restringida de las conclusiones que no serán extensibles a otros niveles de VI.

DISEÑOS FACTORIALES COMPLETAMENTE AL AZAR

Los diseños bivalentes y multivalentes constan sólo de una variable independiente con diferentes niveles, cada uno de los cuales se aplica a distinto grupo de sujetos. Dado que el investigador sólo está interesado en la acción de una variable independiente, se denominan *diseños simples* o *diseños unifactoriales.* En estos diseños se intenta mantener todo constante excepto la variable independiente, con el fin de poder atribuir a dicha variable los cambios observados en la variable dependiente. Sin embargo, se plantean dos dificultades:

1) Es casi imposible mantener constantes todas las variables excepto una. El control completo es muy difícil de conseguir.

2) En la realidad educativa es apropiado hablar de causación múltiple, pues en cada experimento una gran multiplicidad de variables afectan conjuntamente a la variable dependiente.

En consecuencia, más que tratar de hacer lo imposible para mantener constantes todas las variables extrañas (situación artificial y poco generalizable), podemos considerar la acción conjunta de dos o más variables independientes.

En un diseño multifactorial o «factorial», el investigador manipula dos o más variables independientes (cada una de ellas con dos o más niveles de tratamiento) con el fin de estudiar los cambios operados en la variable dependiente como consecuencia de dicha manipulación.

Grupos	*n*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	4	Azar	-	a_1b_1	$\bar{X}_1$
2	4	Azar	-	a_1b_2	$\bar{X}_2$
3	4	Azar	-	a_2b_1	$\bar{X}_3$
4	4	Azar	-	a_2b_2	$\bar{X}_4$

Las respectivas medias de cada grupo se representarían del siguiente modo:

A
Método

B Gratificación		Pasivo	Activo
	Baja	$\bar{X}_1$	$\bar{X}_3$
	Alta	$\bar{X}_2$	$\bar{X}_4$

Las combinaciones de tratamientos, tal como han sido simbolizadas más arriba, son las siguientes:

A
Método

B Gratificación		Pasivo	Activo
	Baja	Grupo 1	Grupo 3
	Alta	Grupo 2	Grupo 4

Grupo 1: Las dos variables actúan con los niveles más bajos, es decir, método pasivo o basado en una participación mínima del alumno y en la poca adjudicación de materiales escolares como recompensa (gratificación baja).
Grupo 2: Los sujetos se someten al nivel más bajo de participación (método pasivo) mientras que la gratificación es máxima.
Grupo 3: Se concede a los alumnos el mayor grado de iniciativa y participación en las clases (método activo), pero el nivel de gratificación es el más bajo.
Grupo 4: Se combinan los niveles más altos de cada factor o variable: método activo y gratificación alta.

EFECTOS FACTORIALES

Los diseños factoriales aportan una información que no podría obtenerse en un diseño unifactorial de una sola variable independiente aunque ésta tenga más de dos niveles de tratamiento. Con un diseño factorial pueden analizarse los siguientes efectos:

a) *Efectos simples.* Expresan los efectos de los niveles de una variable independiente o factor dentro de cada nivel del otro factor. Así, en el ejemplo propuesto, el factor A (método) ejerce dos efectos simples. El primer efecto simple de A se obtiene al comparar las medias de A en el primer nivel del factor B (gratificación baja): $X3-X_¡$, El segundo efecto simple de A se obtiene al comparar las medias de A en el segundo nivel del factor B (gratificación alta): X_4-X_2. De la misma manera se pueden obtener los efectos simples del factor B (gratificación).

b) *Efectos principales o globales.* Se refieren a la influencia global de cada variable independiente por separado, es decir, sin tener en cuenta los valores que toma la otra variable independiente. El efecto principal equivale a la influencia ejercida cuando en el experimento interviene una sola variable independiente. En consecuencia, los efectos principales se pueden estimar calculando las diferencias entre las medias de las filas y entre las medias de las columnas. Ahora bien, dado que los efectos simples son estimaciones del efecto principal, para calcular el efecto principal también pueden promediarse los efectos simples.

c) *Efectos secundarios o de interacción.* Existe interacción entre dos variables independientes cuando el valor que toma la variable dependiente, como resultado de las modificaciones introducidas en aquéllas, depende de los valores que toma la otra variable independiente. La interacción, por tanto, supone que un tratamiento tiene diferentes efectos al combinado con diferentes niveles de otro tratamiento. En definitiva, la influencia simultánea de varios factores hace que ciertas combinaciones entre ellos sean más efectivas que otras.

En el ejemplo propuesto, gracias al diseño factorial podemos comprobar si el efecto del método activo (o del método pasivo) varía a medida que se aplica combinado con los respectivos niveles de la variable gratificación (alta y baja). Si la variación fuera efectiva (significativa) es que existe interacción entre las dos variables o factores (método y gratificación).

La interacción es significativa cuando la diferencia entre los efectos simples de un factor es demasiado grande para ser atribuible al azar. Así, la comparación de los efectos simples para un determinado factor es una estimación de la posible existencia de interacción. En un diseño factorial 2 x 2 la interacción también puede estimarse calculando la diferencia entre la suma de los valores de las diagonales en la tabla de datos promedio.

Para determinar si existe interacción entre dos o más variables puede procederse de dos maneras:

1) *Análisis de la varianza.* Proporciona información sobre la significación estadística de la interacción. Su estudio pormenorizado se expondrá en capítulos posteriores al tratar el análisis de datos.

2) Representación gráfica. La inspección visual de la gráfica permite apreciar de forma intuitiva si existe o no interacción. En el eje de abscisas se representan los valores de uno de los factores y en el de ordenadas se representan los valores de la variable dependiente. Al unir los puntos correspondientes a los niveles del segundo factor, las rectas resultantes pueden tomar alguna de las siguientes formas:

- rectas paralelas;
- rectas que no son paralelas, aunque no llegan a cortarse en la gráfica;
- rectas que se cortan en la gráfica.

Ilustraremos la interpretación gráfica de la interacción a partir del ejemplo ya planteado para comprobar cómo influyen en el rendimiento el método didáctico utilizado (pasivo-activo) y la gratificación (alta-baja). También se describirán las tres situaciones contemplando distintos tipos de efectos factoriales como son: los efectos principales o globales, secundarios o de interacción, y simples.

PRIMERA SITUACIÓN: RECTAS PARALELAS

Supongamos que en el ejemplo planteado las puntuaciones de rendimiento obtenidas por los 16 sujetos en el postest fueran las siguientes:

Sujeto	*Tratamiento*	*Rendimiento*
1	a_1b_1	0
2	a_1b_1	1
3	a_1b_1	2
4	a_1b_1	1
5	a_1b_2	3
6	a_1b_2	4
7	a_1b_2	3
8	a_1b_2	2
9	a_2b_1	4
10	a_2b_1	5
11	a_2b_1	6
12	a_2b_1	5
13	a_2b_2	8
14	a_2b_2	6
15	a_2b_2	8
16	a_2b_2	6

Estos datos pueden tabularse de distintas maneras. Cuando queremos analizar los efectos globales de cada variable puede ser útil una tabla como la siguiente:

		A		
		a_1	a_2	$\bar{X}$
B	b_1	0 1 2 1	4 5 6 5	3
	b_2	3 4 3 2	8 6 8 6	5
	$\bar{X}$	2	6	4

Como ejemplo ilustrativo se han registrado todas las puntuaciones directas de rendimiento; sin embargo, presenta mayor utilidad incluir las medias obtenidas por cada grupo en una tabla de datos promedio como se verá más adelante.

Para apreciar con más claridad el efecto del método en cada nivel de gratificación también puede elaborarse una tabla como la siguiente:

Método	*Pasivo*		*Activo*	
Gratificación	Baja	Alta	Baja	Alta
Puntuaciones	1	4	5	6
de rendimiento	2	3	6	8
	1	2	5	6
Rendimiento medio	1	3	5	7

A partir de las tablas anteriores pueden estimarse los efectos factoriales principales o globales, secundarios o interacciones, y simples o diferenciadores.

a) *Efectos factoriales principales o globales*

Se refieren a la influencia global de las variables método y gratificación consideradas por separado. En definitiva, es como si hubiéramos realizado dos experimentos bivalentes, uno para ver qué método es más efectivo (el activo o el pasivo) y otro para ver qué gratificación es más efectiva (la alta o la baja). La comparación de los resultados finales de cada columna nos permite comprobar si los sujetos rinden más con un método pasivo que con un método activo. Al comparar los resultados de las filas podemos estimar la eficacia relativa de cada tipo de gratificación (baja o alta):

1) Efecto global de la gratificación. El efecto principal de la gratificación vale 5 - 3 = 2. Es decir, al pasar de un nivel de gratificación bajo a un nivel alto el rendimiento aumenta en 2 unidades:

Gratificación	*Baja*	*Alta*
	0	3
	1	4
	2	3
Puntuaciones	1	2
de rendimiento	4	8
	5	6
	6	8
	5	6
Rendimiento medio	3	5

2) Efecto global del método. El efecto principal del método vale 6- 2 = 4. Es decir, al pasar de un método pasivo a un método activo el rendimiento aumenta en cuatro unidades:

Método	*Pasivo*	*Activo*
	0	4
	1	5
	2	6
Puntuaciones	1	5
de rendimiento	3	8
	4	6
	3	8
	2	6
Rendimiento medio	2	6

b) *Efecto factorial secundario o interacción*

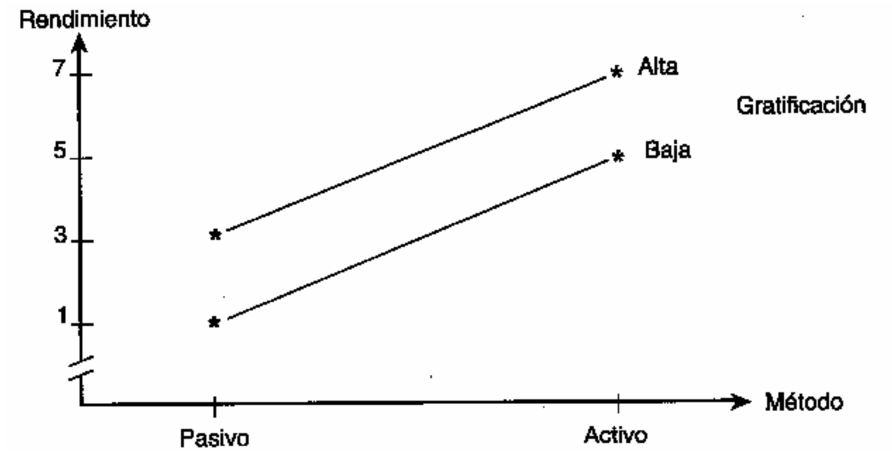
Las conclusiones anteriores de que la gratificación alta es más efectiva que la baja porque aumenta el rendimiento en dos unidades (suponiendo que hubiera significación estadística) y de que el método activo es más efectivo que el pasivo tienen sentido en el caso de que no existan efectos factoriales secundarios o interacciones.

No existe interacción entre la gratificación y el método cuando el aumento del rendimiento provocado por la gratificación alta (dos unidades) es el mismo al combinarse con el método pasivo y con el activo. Es decir, al pasar de una gratificación baja a otra alta el rendimiento aumenta una cantidad constante que siempre vale dos unidades, tanto si el método es pasivo como si es activo. Esto significa que los efectos simples correspondientes a los niveles de una misma variable o factor son estadísticamente iguales. La tabla de datos promedio es la siguiente:

Método

Gratificación	Pasivo	Activo	
Baja	1	5	3
Alta	3	7	5
	2	6	4

Puede comprobarse que al pasar de un método pasivo a un método activo el rendimiento también aumenta una cantidad constante de cuatro unidades, tanto para la gratificación baja como para la gratificación alta. Es decir, los efectos simples vuelven a ser iguales. La ausencia de interacción o igualdad de efectos simples queda reflejada por la presencia de dos rectas paralelas:



La interacción puede estimarse a partir de la diferencia entre la suma de los valores de las diagonales:

$$[(1 + 7) - (3 + 5)] /2 = O.$$

Cuando la interacción no es significativa se puede concluir que los efectos factoriales son «aditivos» o que existe aditividad, y por tanto, que existe independencia entre los factores o variables. Aunque la interacción no sea significativa, en algunos casos el análisis de los efectos simples puede ser muy revelador.

c) *Efectos factoriales simples*

Si nos fijamos en la variable gratificación, podemos determinar dos efectos de la misma:

1) Efecto simple de la gratificación cuando el método es pasivo: 3- 1 = 2.
2) Efecto simple de la gratificación cuando el método es activo: 7- 5 = 2.

Esta información indica que el efecto de la gratificación tiene la misma intensidad al combinarse con ambos métodos. Con respecto a la variable método puede procederse de la misma manera:

3) Efecto simple del método cuando la gratificación es alta: 7- 3 = 4.
4) Efecto simple del método cuando la gratificación es baja: 5- 1 = 4.

Estos resultados indican que el método tiene una influencia idéntica cuando la gratificación es baja y cuando es alta. Cuando las diferencias entre los efectos factoriales simples de una variable pueden atribuirse estadísticamente al azar no existen efectos factoriales secundarios o interacción. Así, para la variable gratificación, sus efectos simples (2 y 2) son dos estimaciones de la influencia de dicha variable. Si son estadísticamente iguales, hay que concluir que la variable método no interviene en ellos y que no existe interacción. Es decir, la variable gratificación, al combinarse con los dos tipos de método, influye por igual en el rendimiento.

Puede comprobarse que el efecto principal o influencia global de cada variable puede estimarse también a partir de los efectos simples:

- Influencia global de la gratificación: promedio de los efectos simples (2 + 2) / 2 = 2. Como ya hemos visto, dicho efecto se puede calcular recurriendo a la diferencia entre las medias de las filas: 5- 3 = 2.
- Influencia global del método: promedio de los efectos simples (4 + 4) / 2 = 4.

El mismo efecto se puede calcular con la diferencia entre las medias de las columnas: 6-2 = 4.

También puede estimarse la interacción a partir de la diferencia entre los efectos simples, ya sea 2- 2 = 0, o bien, 4- 4 = 0.

SEGUNDA SITUACIÓN: LAS RECTAS NO SON PARALELAS PERO NO SE CORTAN EN LA GRÁFICA

También podríamos haber obtenido datos ficticios como los siguientes:

		Método		
		Pasivo	Activo	
Gratificación	Baja	3	1	2
	Alta	5	7	6
		4	4	4

Según la tabla, los tipos de efectos son:

a) *Efectos factoriales principales* o *globales*

1) Influencia global de la gratificación: 6 - 2 = 4. Estadísticamente podría existir diferencia significativa y ser más efectivo el nivel alto de gratificación.

2) Influencia global del método: 4- 4 = O. El método activo tiene la misma efectividad que el método pasivo, ambos tienen un efecto idéntico.

b) *Efecto factorial secundario* o *interacción*

Las conclusiones anteriores son globales y pueden inducir a error. No son válidas por igual cuando cada variable se combina con los niveles de la otra. Según los efectos globales de la gratificación, al pasar de un nivel bajo de gratificación a un nivel alto aparece un aumento en rendimiento de cuatro unidades. Si este aumento fuera estadísticamente igual (cuatro unidades) bajo los dos métodos (activo y pasivo), no existiría interacción y el efecto global sería válido. Sin embargo, bajo el método pasivo el aumento es de sólo dos unidades, pero bajo el método activo el aumento es de seis unidades, lo que origina una aumento global medio de (2 + 6) / 2 = 4 unidades, como ya se había calculado. Este aumento es debido al cambio introducido en los niveles de gratificación.

Si sólo nos basamos en los efectos globales del método, el error puede ser aún más grave. Según los efectos principales o globales, tanto el método pasivo como el activo tienen la misma efectividad (4 - 4 = O). Ahora bien, cuando la gratificación es baja, al pasar de un método pasivo a un método activo el rendimiento sufre un descenso de dos unidades (1- 3 = -2). En cambio, cuando la gratificación es alta, aparece un aumento de 7 - 5 = 2 unidades; de ahí que se detecte un aumento global medio que vale

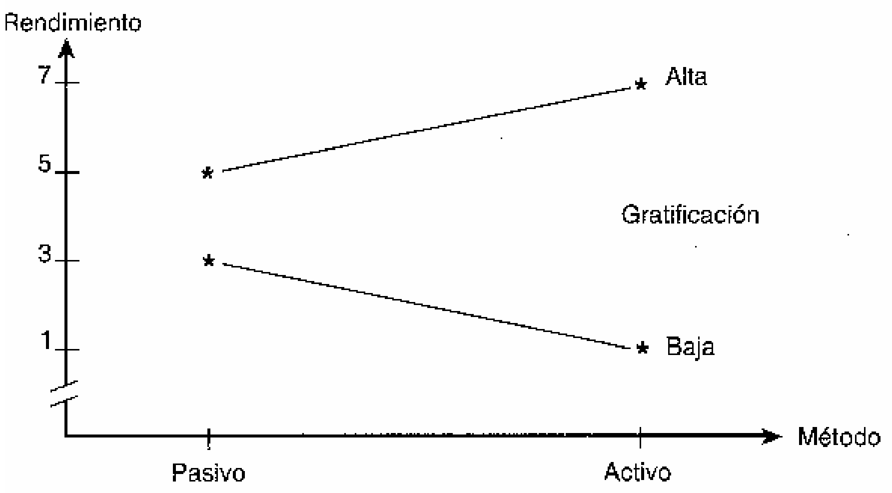
[2 + (- 2)] /2 = O.

Es decir, a pesar de que ambos métodos tienen el mismo efecto global, sus efectos factoriales simples son diferentes para cada nivel de gratificación.

Se produce una interacción entre las variables método y gratificación cuando los cambios observados en el rendimiento como consecuencia del método son significativamente diferentes para cada tipo de gratificación; es decir, la efectividad del método es distinta según el nivel de gratificación, y viceversa. El cálculo numérico de la interacción sería

$$[(3 + 7) - (5 + 1)] /2= 2$$

Al existir interacción no es suficiente analizar los efectos globales o principales de cada variable; es imprescindible analizar los efectos simples. Cuando existe interacción las rectas no tienden a ser paralelas:



c) *Efectos factoriales simples*

Con respecto a la variable gratificación hemos estimado dos efectos de la misma:

1) Efecto factorial simple de la gratificación cuando el método es pasivo: 5-3 = 2.
2) Efecto factorial simple de la gratificación cuando el método es activo: 7-1 = 6.

En consecuencia, el efecto de la gratificación tiene diferente intensidad (debería comprobarse si la diferencia es excesiva para ser atribuida sólo al azar) cuando el método es pasivo y cuando es activo.

Con respecto a la variable método puede procederse de la misma manera:

3) Efecto factorial simple del método cuando la gratificación es alta: 7-5 = 2.
4) Efecto factorial simple del método cuando la gratificación es baja: 1-3 = -2.

Como puede apreciarse, el efecto es negativo porque el nivel más participativo del método (activo) es menos efectivo que el nivel más bajo de participación (pasivo).

Estos resultados indican que el método tiene una influencia distinta según sea el nivel de gratificación (alta o baja). Cuando la gratificación es alta, al pasar de un método pasivo a un método activo se aprecia que el rendimiento aumenta en dos unidades. En cambio, cuando la gratificación es baja, ocurre al contrario: al pasar de un método pasivo a un método activo el rendimiento disminuye en dos unidades. Como es evidente, los efectos globales pueden calcularse recurriendo a los efectos globales:

- Influencia global de la gratificación: promedio de los efectos simples (2 + 6) *12* = 4. El mismo efecto se puede calcular a partir de la diferencia entre las medias de las filas: 6 - 2 = 4. - Influencia global del método: promedio de los efectos simples [2 + (- 2)] *12* = 0.

De la misma manera, a partir de la diferencia entre las medias de las columnas: 4 - 4=0.

TERCERA SITUACIÓN: LAS RECTAS SE CORTAN EN LA GRÁFICA

Otra posibilidad es que los datos resultantes sean como los de la tabla que se adjunta:

a) *Efectos factoriales principales o globales*

1) Influencia global de la gratificación: 5 - 5 = 0. Los dos niveles de gratificación tienen la misma efectividad.

2) Influencia global del método: 6- 4 = 2. El método activo aumenta dos unidades el rendimiento con respecto al método pasivo.

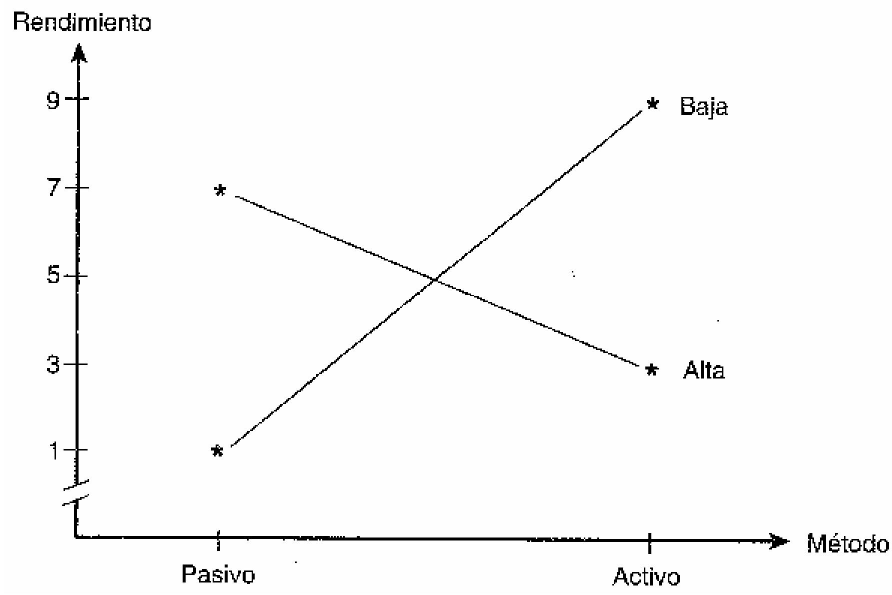
Método

		Pasivo	Activo	
Gratificación	Baja	1	9	5
	Alta	7	3	5
		4	6	5

b) *Efecto factorial secundario a interacción*

Cuando en la gráfica aparecen rectas que se cortan es muy posible que exista interacción entre los dos factores (gratificación y método), cuyo valor sería:

[(1 + 3) - (7 + 9)] *12* = - 6



Cuando exista interacción es necesario analizar los efectos simples.

c) *Efectos factoriales simples*

1) Efecto factorial simple de la gratificación cuando el método es pasivo: 7 - 1 = 6.
2) Efecto factorial simple de la gratificación cuando el método es activo: 3 - 9 = - 6.

Con respecto a la variable método puede procederse de la misma manera:

3) Efecto factorial simple del método cuando la gratificación es alta: 3 - 7 = - 4.
4) Efecto factorial simple del método cuando la gratificación es baja: 9 - 1 = 8.

Para comprobar los efectos globales:

- Influencia global de la gratificación: promedio de los efectos simples:

[6 + (- 6)] / 2 = 0, o bien 5 - 5 = 0.

- Efecto global del método: [8 + (- 4)] /2 = 2, o bien, 6 - 4 = 2.

5.7 DISEÑOS DE BLOQUES HOMOGÉNEOS AL AZAR

En los diseños completamente al azar la heterogeneidad entre los sujetos puede originar una varianza intragrupo (o del error) muy elevada. Con el fin de reducir la variabilidad de los sujetos y hacer más sensible el experimento se puede introducir la

técnica del bloqueo. De esta manera se pueden apreciar mejor los efectos de VI. Supongamos que estamos interesados en ver el efecto de tres niveles de gratificación (A = alto, M = medio y B = bajo) en el rendimiento de 12 niños de primaria. Para realizar un experimento de bloques hay que tener en cuenta las siguientes fases:

1) Mídase alguna variable extraña que sospechemos que pueda estar estrechamente relacionada con la variable dependiente. En consecuencia, podría contaminar el experimento. Dicha variable se denomina *variable de bloqueo* o, en su caso, *de apareo* (cuando cada bloque está formado por dos sujetos). Generalmente, dicha variable no coincide con la variable dependiente. En el ejemplo planteado consideramos que la variable inteligencia sería una variable de bloqueo adecuada. Después de medir la inteligencia obtenemos las siguientes puntuaciones:

Sujetos	*Inteligencia*
Antonio	98
Rosa	100
Carlos	94
Ignacio	94
Gema	96
Maribel	98

Sujetos	*Inteligencia*
Mónica	96
José	100
Laura	94
Ana	98
Nuria	100
Raquel	96

2) En función de las puntuaciones obtenidas en la variable de bloqueo, fórmense grupos de sujetos (bloques) con puntuaciones iguales o lo más similares posible. Dentro de cada bloque se asignan al azar, como mínimo, tantos sujetos o unidades experimentales como niveles tenga la variable independiente:

Bloques	*CI*		*Sujetos*	
Bloque 1	94	Ignacio	Carlos	Laura
Bloque 2	96	Gema	Mónica	Raquel
Bloque 3	98	Maribel	Antonio	Ana
Bloque 4	100	Nuria	José	Rosa

3) Asígnese cada nivel de la variable independiente al azar:

			Gratificación	
Bloques	*CI*	*Baja*	*Media*	*Alta*
Bloque 1	94	Ignacio	Carlos	Laura
Bloque 2	96	Gema	Mónica	Raquel
Bloque 3	98	Maribel	Antonio	Ana
Bloque 4	100	Nuria	José	Rosa

4) Aplíquense los niveles de la variable independiente.
5) Mídase la variable dependiente:

		Gratificación		
Bloques	el	*Baja*	*Media*	*Alta*
Bloque 1	94	9	9	9
Bloque 2	96	8	7	6
Bloque 3	98	6	5	4
Bloque 4	100	4	4	2

Si cada bloque está formado por dos sujetos, puede tratarse de un diseño de grupos apareados. Cuando cada unidad experimental está formada por grupos de sujetos, la variable de bloqueo puede considerarse como otra variable independiente o factor, y el experimento se convierte en factorial.

El diseño planteado se denomina *de bloques al azar* 3 4 porque el factor manipulado adopta tres niveles y se han formado cuatro bloques de sujetos con la variable de bloqueo. La representación simbólica del diseño sería la siguiente:

Grupos	*n*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	3	BA_4	-	a_1	$\bar{Y}_1$
2	3	BA_4	-	a_2	$\bar{Y}_2$
3	3	BA_4	-	a_3	$\bar{Y}_3$
4	3	BA_4	-	a_4	$\bar{Y}_4$

La técnica del bloqueo adolece de tres limitaciones importantes. En primer lugar, cuando se ha elegido un sujeto o grupo de sujetos para formar un bloque, el investigador debe seleccionar más sujetos semejantes entre los restantes para completar el bloque iniciado. Con frecuencia es difícil encontrar sujetos suficientes. Esta circunstancia origina la segunda limitación: los sujetos que no son similares a los incluidos en los bloques deben excluirse, con lo que disminuye la representatividad de la muestra con respecto a la población. Por último, la homogeneidad entre los sujetos de un mismo bloque se establece con respecto a la variable de bloqueo, pero los sujetos incluidos en un mismo bloque pueden diferir en alguna otra variable.

5.8 DISEÑOS INTRAGRUPO O DE MEDIDAS REPETIDAS

La eficacia y sensibilidad de un experimento aumenta en proporción a la homogeneidad de los sujetos, ya que, con grupos experimentales más homogéneos, se tiene mayor probabilidad de comprobar, en el caso de que exista, la posible influencia

de la variable independiente sobre la variable dependiente. La varianza del error disminuye progresivamente al pasar de los diseños completamente al azar a los diseños de bloques homogéneos. Con los diseños intragrupo la técnica del bloqueo se lleva hasta sus últimas consecuencias. Si a cada sujeto se le aplican todos los niveles de la variable independiente y se comparan entre sí los efectos producidos, hemos llevado hasta el extremo las condiciones de homogeneidad con respecto a las diferencias individuales. De hecho, cada sujeto puede considerarse como un solo bloque.

En los diseños intragrupo es importante el orden de aplicación. Los resultados obtenidos en la variable dependiente pueden depender de la secuencia seguida en la aplicación de los tratamientos. Para evitarlo es viable aleatorizar el orden de aplicación o puede procederse a un contrabalanceo de los tratamientos.

La representación simbólica de un diseño intragrupo puede realizarse del siguiente modo:

Grupos	*número de sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n	A	-	$a_1; a_2; a_3$	$\bar{X}_1; \bar{X}_2; \bar{X}_3$

Veremos dos aplicaciones de los diseños intragrupo: el diseño tratamiento por sujetos y el análisis de tendencias.

DISEÑO DE TRATAMIENTOS POR SUJETOS

Cuando estamos interesados en comparar la efectividad de una serie de tratamientos o niveles de la variable independiente se utiliza el denominado *diseño de tratamientos por sujetos.*

La representación simbólica anterior puede hacerse extensible del siguiente modo:

Grupos	*Pretest (VI)*	*Tratamiento (VD)*	*Postest*
1	-	$a_1\ a_2\ a_3 \ldots a_k$	$X_1\ X_2\ X_3 \ldots X_k$
1	-	$a_1\ a_2\ a_3 \ldots a_k$	$X_1\ X_2\ X_3 \ldots X_k$
1	-	$a_1\ a_2\ a_3 \ldots a_k$	$X_1\ X_2\ X_3 \ldots X_k$
•			
•			
•			
n	-	$a_1\ a_2\ a_3 \ldots a_k$	$X_1\ X_2\ X_3 \ldots X_k$
			$\bar{X}_1\ \bar{X}_2\ \bar{X}_3 \ldots \bar{X}_k$

En la medida de lo posible, el orden de aplicación de los niveles de la variable independiente será totalmente aleatorio. La estructura lógica de este diseño es similar a la del diseño de bloques, ya que cada sujeto equivale a un bloque.

ANÁLISIS DE TENDENCIAS

Se utiliza cuando el énfasis se centra en la mejora o cambio que se produce como consecuencia de una práctica más o menos prolongada. Los niveles de la variable independiente difieren entre sí según una cantidad constante. Se trata de comprobar si las medias obtenidas en la variable dependiente bajo cada tratamiento tienden a variar en un sentido lineal, curvilíneo, etc., y hasta qué punto es significativa esta tendencia. Es un diseño muy común en situaciones de aprendizaje donde interesan los efectos acumulados de los tratamientos (fig. 5.5).

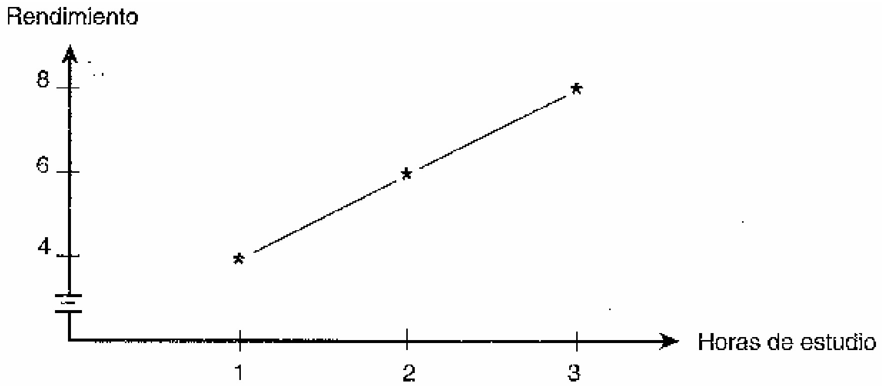


FIG. 5.5 Función lineal de las medias de rendimiento según las horas de estudio en un experimento intragrupo

5.9 TIPOS DE EXPERIMENTOS

Para conseguir el máximo grado de control posible en los experimentos se tiende a un aislamiento del contexto natural donde se produce el fenómeno. La finalidad de este aislamiento radica en el objetivo de eliminar las múltiples influencias extrañas capaces de actuar juntamente con la variable independiente, pues la acción de dichas variables la confundiríamos con la influencia de la variable independiente sobre la variable dependiente.

En la medida que controlamos las variables extrañas (restricciones) garantizamos la validez interna, pero las condiciones en que se desarrolla el experimento son más artificiales y, por tanto, menos representativas, pues en un contexto real las variables no se hallan tan controladas. En consecuencia, la generalización a situaciones reales queda reducida. De ahí la conveniencia de diferenciar entre experimentos de laboratorio y de campo, teniendo en cuenta que entre ambas modalidades sólo existe una diferencia de grado, los primeros con un control máximo, y los segundos, con un control menor.

Experimento de laboratorio. Es un estudio de investigación donde se consigue un aislamiento del contexto natural en que se produce el fenómeno con el fin de eliminar las múltiples influencias extrañas capaces de afectar a la variable dependiente.

Experimento de campo. Es un estudio de investigación en una situación real donde una o más variables independientes son manipuladas por el experimentador en condiciones controladas con el máximo rigor que permita la situación.

5.10 POSIBILIDADES Y LÍMITES

El método experimental se aplica a través de diseños específicos que suponen adoptar variaciones respecto a la categoría global, y que posibilitan responder de forma más acorde a determinados problemas y situaciones de investigación. En cuanto a sus aplicaciones a cuestiones o problemas educativos, los diseños más rigurosos se han utilizado principalmente en el área de tecnología educativa y didáctica, configurando la inmensa mayoría de trabajos que se han publicado.

a) *Posibilidades.* Según algunos autores (Mayntz, 1983), es el más prominente de todos los métodos de investigación en cuanto que es el único que permite constatar estrictamente relaciones causales. Al manipular o provocar el fenómeno, el investigador está en condiciones idóneas para inferir proposiciones causales (Cook y Campbell, 1979; Arnau, 1978; Haimson y Elfebein, 1985). Otras funciones básicas del método experimental son descubrir relaciones en condiciones puras y no contaminadas, probar predicciones derivadas de la teoría o de otra investigación y desarrollar teorías e hipótesis para elaborar sistemas teóricos.

Estas funciones son viables gracias al elevado grado de control que se ejerce sobre:

- Las variables independientes, mediante su modificación en cantidad, grado o cualidad.
- Las variables extrañas, cuyo dominio o control facilitará la maximización de la varianza primaria.
- La medición, que deberá ser fiable, precisa y válida.

El control relativamente completo permite una gran especificidad en las definiciones operativas de las variables, con lo que se favorece la precisión y la posibilidad de llevar a cabo las réplicas necesarias. La investigación queda muy bien delimitada, es inequívoca y los errores de medida son mínimos, ya que las condiciones del experimento se especifican minuciosamente. Como consecuencia, la validez interna es alta, con lo que estamos ante un enfoque adecuado para probar diferentes aspectos de la teoría.

Otra característica vinculada a la necesidad de inferir efectos de los tratamientos es la utilización de la comparación. Existen muchos tipos de comparación en la experimentación que responden a diferentes propósitos y originan distintas modalidades de diseños de investigación. También cabe destacar su precisión y sensibilidad a la hora de maximizar la varianza experimental y minimizar la del error, como consecuencia del control de las situaciones en las que se produce el experimento (Kerlinger, 1982). En resumen, las ventajas y posibilidades de esta metodología se pueden resumir en precisión y economía (Mouly, 1978).

Los *experimentos de laboratorio* son especialmente idóneos para cubrir objetivos de investigación como los siguientes:

- Constatar relaciones en condiciones «puras» e incontaminadas.
- Someter a prueba las predicciones derivadas de la teoría (hipótesis).
- Generar nuevas investigaciones.
- Refinar las teorías y las hipótesis. Ayudan a construir sistemas teóricos.

El experimento de laboratorio presenta ventajas como las siguientes:

- Control relativamente completo.
- Permite la asignación aleatoria y la manipulación de una o más variables independientes.
- Alto grado de especificidad en las definiciones operativas de las variables.
- Precisión. Bien delimitado e inequívoco. La varianza del error es pequeña, ya que las condiciones del experimento se especifican minuciosamente disminuyendo los errores.
- Alta validez interna.
- Flexibilidad: hay muchas posibilidades experimentales y se pueden poner a prueba variados aspectos de la teoría.
- Replicabilidad: posibilidad de repetir el experimento introduciendo pequeñas variantes.

Las ventajas principales de los experimentos de campo pueden ser las siguientes:

- Muy utilizados en pedagogía, se llevan a cabo en aulas, colegios y otras instituciones y comunidades. Las variables actúan en su propio medio natural. - Al menos en teoría permiten manipular las VI y asignar sujetos y tratamientos al azar.
- Las variables suelen tener un efecto más potente que las de laboratorio. Cuanto más realista es la situación más potentes serán las variables.
- Adecuados para estudiar las influencias sociales complejas, los procesos y cambios sociales en situaciones de la vida real.
- Pueden utilizarse para someter a prueba las teorías y para solucionar problemas prácticos.
- Flexibilidad y aplicabilidad a una gran variedad de problemas.

b) *Límites.* Pero los experimentos no están exentos de algunos inconvenientes.

Como contrapartida existen limitaciones al aplicar esta metodología al campo pedagógico, como son la selección de las variables, la naturaleza re activa de los sujetos, posibles sesgo s introducidos por el experimentador, limitaciones procedentes del muestreo, problemas deontológicos (Pérez Juste, *1985a;* Mialaret, 1972), además de los de generalización como consecuencia de la artificialidad del control (Cook y Campbell, 1979).

Así, dada la artificialidad de la situación experimental, puede ocurrir que el influjo de la variable independiente sobre la dependiente sea débil, ya que la situación se ha provocado con propósitos especiales. Por ello, hay que tener en cuenta que los experimentos, especialmente los de laboratorio, presentan poca generalización. Esta circunstancia disminuye su validez externa, con el consiguiente riesgo de interpretar o extrapolar erróneamente los resultados.

Algunas de las limitaciones comentadas deben matizarse si se tiene en cuenta que los experimentos, más que generalizar, pretenden constatar relaciones de causalidad en condiciones puras y no contaminadas, sometiendo a prueba las predicciones derivadas de la teoría. De ahí que **un** cometido fundamental del enfoque experimental sea refinar las teorías a través del contraste de hipótesis, con lo que se favorece la construcción de sistemas teóricos.

Aunque parece ser el método idóneo para garantizar la validez interna, algunos autores (Harsé y Secord, 1976) la cuestionan a la hora de estudiar la conducta social, señalando limitaciones como las que impone la misma naturaleza de la situación, la operacionalización de conceptos, el desconocimiento de parámetros relevantes y el tipo de interacción peculiar de los experimentos.

Algunas limitaciones de los experimentos de laboratorio son las siguientes:

- Debilidad de las variables independientes, ya que se han creado las situaciones con propósitos especiales.
- Artificialidad de la situación experimental si no se tienen en cuenta los objetivos de los experimentos de laboratorio (relaciones de causación entre variables).
- Baja validez externa, poca generalización, pero su objetivo primordial no es generalizar.

En cuanto a los experimentos de campo, pueden destacarse las siguientes limitaciones:

- Carecen de **un** control riguroso.
- Objeciones ante la manipulación de las variables que puedan afectar a la educación o al rendimiento del alumno.
- Falta de precisión.
- Requieren en cierta medida **un** experto en trabajo social.

CAPÍTULO 5

Metodología experimental

El proceso general de la investigación ya se ha descrito en el capítulo 3; sin embargo, cada método presenta variaciones que le permiten adaptarse a las exigencias de problemas específicos de investigación. En consecuencia, es apropiado apuntar algunas características inherentes al método experimental, abordando posteriormente aspectos como la varianza y las técnicas para controlarla, las fuentes de invalidez y la estrategia que hay que seguir en los diseños más importantes. A su vez, se contemplan las posibilidades y límites de dichos diseños y se ofrecen orientaciones para planificar un experimento.

Los estudios experimentales estrictos se caracterizan por presentar los rasgos siguientes:

1) El investigador manipula la variable independiente y le asigna niveles.

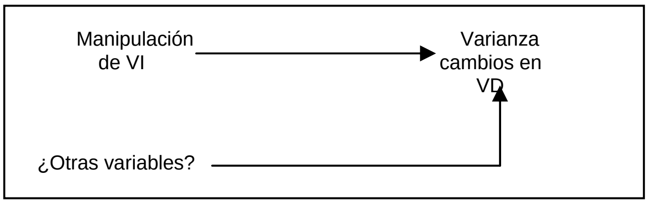
2) Han de aplicarse dos o más niveles de la variable independiente con el fin de poder contrastar los efectos sobre la variable dependiente. Estos efectos se estiman comparando los resultados obtenidos en la variable dependiente.

3) La muestra se elegirá al azar. Si los niveles de la variable independiente se aplican a grupos de sujetos distintos, antes de comenzar el experimento los sujetos se asignan al azar a los grupos donde se aplicarán los niveles de la variable independiente. A cada grupo se le aplica al azar un nivel de la variable independiente. Así se garantiza que por efecto del azar los grupos sean homogéneos o equivalentes, para que después de aplicar los niveles de la variable independiente dichos grupos sean comparables entre sí. De esta manera, las posibles diferencias encontradas al medir la variable dependiente en cada grupo son atribuibles a la aplicación de los niveles de la variable independiente.

5.1 TIPOS DE VARIANZA

La meta ideal del experimentador es que los cambios o variaciones en los valores de la variable dependiente (VD) sean atribuibles a las manipulaciones de la variable independiente (VI) y no a otras variables.

Como ya se ha expuesto, estas variables extrañas podrían influir con intensidad



desigual en los grupos control y experimental(es) y sus efectos se unirían al de la variable independiente, constituyendo fuentes de variación. En consecuencia, el investigador no podría distinguir qué cambios en la variable dependiente son exclusivamente debidos a la variable independiente, y el experimento quedaría invalidado, a menos que dichas variables se controlen igualando o eliminando su influencia. El objetivo del investigador es distinguir los cambios de la variable dependiente que se deben a la manipulación de la variable independiente y los que se deben a otras variables (extrañas). Tanto la variable independiente como las otras variables son fuentes de variación o varianza. Analizaremos las más importantes agrupadas en tres grandes bloques o fuentes de variación:

a) *El experimentador.* Las actitudes y expectativas del investigador sobre los efectos de los niveles de la variable independiente pueden influir en los resultados del experimento. También los rasgos físicos (edad, sexo, aspecto, etc.), la forma de comportarse y las características sociales (status, nivel de estudios, etc.) del experimentador pueden influir.

b) *El contexto experimental.* La posible contaminación de los datos en función del contexto puede provenir de dos fuentes: de cómo captan los sujetos las condiciones de la situación experimental y del diseño (nivel de las variables ambientales, condiciones de aplicación, etc). Veamos brevemente estos aspectos:

- *Demanda de la situación experimental.* La interpretación que los sujetos de los grupos experimentales realizan acerca del experimento puede condicionar su actuación. A través de esta interpretación el sujeto intuye lo que se espera de él (demanda de la situación experimental), y lo puede descubrir durante la realización del experimento, a través de la captación de indicios, y antes del experimento por medio de los rumores sobre la investigación, el lugar donde se realiza y quiénes lo llevan a cabo.

- *Relacionadas con el diseño.* Pueden provenir del medio ambiente (ruido, iluminación, temperatura, etc.) o de las condiciones de aplicación (instrucciones, sistemas de medición, presentación de materiales, estímulos y recogida de datos).

c) *El sujeto experimental.* Cuando los experimentos se llevan a cabo con sujetos universitarios, obligados o voluntarios, hay que cuestionar su representatividad con respecto al resto de la población. Antes de aplicarse los niveles de la variable independiente los grupos pueden diferir en variables como edad, sexo, motivación,

inteligencia o personalidad. Los sujetos de los grupos experimentales pueden verse más motivados por la novedad que los sujetos que pertenecen al grupo control.

Debido a que la acción de estos tres grupos de variables o fuentes de variación es simultánea, cuando se ha elegido una variable independiente, el investigador ha de controlar o igualar la influencia de las restantes fuentes de variación que puedan afectar a la variable dependiente. Procurará que influyan por igual en el grupo control y en el experimental. Si todas las variables extrañas influyen por igual en ambos grupos, las diferencias (varianza) encontradas entre el grupo control y el experimental serán debidas a la variación introducida deliberadamente (niveles de VI).

Supongamos que queremos comprobar la eficacia de un programa basado en la utilización de medios audiovisuales (MA V) para la enseñanza del inglés en primaria. Para ello, seguimos las fases de la aleatorización, y después de comprobar la equivalencia inicial de los grupos (pretest), aplicamos el programa con MA V a un grupo experimental, y a un grupo control le aplicamos el programa sin MA V. Posteriormente, con una prueba paralela al pretest, medimos el rendimiento final en lengua inglesa (postest) y obtenemos las medias especificadas en la tabla 5.1.

TABLA 5.1 Ejemplo de diseño experimental para estudiar la eficacia de los MA V

Grupo	*n*	*Asignación*	*Pretest*	*Programa*	*Postest*
Experimental (GE)	5	Azar	2	Con MAV	9
Control (GC)	5	Azar	2	Sin MAV	5

Con las medias obtenidas podemos efectuar distintas comparaciones. Imaginemos que después de aplicar las pruebas estadísticas adecuadas obtenemos los resultados que aparecen en la tabla 5.2.

TABLA 5.2 Algunas comparaciones que pueden efectuarse en un diseño experimental

Objetivo	*Prueba estadística*	*Medias*	*Decisión*
Equivalencia inicial	t datos independ.	2 - 2	Ha
Equivalencia final (eficacia MA V)	t datos independ.	9 - 5	Hj
Cambio en GE	t datos apareados	9 - 2	H,
Cambio en GC	t datos apareados	2 - 5	H,

Cuando el investigador pretende contrastar una hipótesis a través de un experimento puede proceder del siguiente modo:

1) Mantiene constantes o iguala las variaciones (varianza) que puedan producir todas las posibles influencias que procedan de las variables extrañas. Esto es, controlar la influencia de las variables extrañas.
2) Introduce una variación deliberada controlando (manipulando) la variable independiente.
3) Analiza la intensidad de la variación que se produce en VD como consecuencia de la variación introducida al aplicar distintos niveles de la variable independiente.

De ahí que las conclusiones de un experimento se basen en las variaciones, cambios o varianzas observadas en la variable dependiente a raíz de las variaciones o cambios introducidos deliberadamente en la variable independiente, procurando igualar o sistematizar la varianza debida a variables extrañas. Para entender mejor la importancia de la varianza la analizaremos más detenidamente en el contexto del experimento planteado sobre los MA V.

En la tabla 5.3 se han esquematizado, de forma ficticia y aproximada, algunas fuentes de variación que podrían intervenir en el experimento de los MA V. Las puntuaciones en inglés obtenidas por los sujetos en el pretest presentan ciertas diferencias o variaciones que pueden atribuirse a tres fuentes fundamentales: diferencias individuales (motivación, inteligencia...), influencias ambientales (profesores, familia...) e imperfecciones de los instrumentos de medida. A pesar de esta varianza observada en las puntuaciones de los sujetos los grupos que se han formado tienen la misma media en inglés. Es decir, puede existir cierto grado de variación (debería constatarse estadísticamente) entre los sujetos que forman cada grupo (varianza intragrupo) sin que aparezcan necesariamente diferencias entre los grupos (varianza intergrupos).

Al iniciar el experimento, el investigador introduce una variación: utiliza los MA V en el grupo experimental, pero no en el grupo control. Ahora bien, como puede apreciarse en la tabla 5.3, otras variables como el profesor y la memoria de los sujetos también pueden originar variaciones.

El posible efecto de todas las fuentes de variación mencionadas queda recogido en la varianza total que presentan las puntuaciones en el postest de todos los sujetos. Como puede apreciarse en la tabla 5.3, la variación total de las puntuaciones de inglés se debe a unas variaciones constantes o sistemáticas que constituyen la denominada *varianza sistemática* y a otras irregulares o aleatorias que constituyen la *varianza residual o del error.*

A) *Varianza sistemática.* Es una variación originada por variables que desvían los datos en una dirección más que en otra, aumentando o disminuyendo, pero siempre en la misma magnitud y dirección. Esta desviación, en el experimento perfectamente controlado, es debida exclusivamente a la variable independiente (MA V). Cuando el grado de control no es el ideal, la desviación ha podido ser provocada también por variables extrañas que no han sido convenientemente controladas. Así, en el ejemplo propuesto el entusiasmo del profesor ha influido más en el grupo experimental (+ 2) que en el de control (+ 1).

En cambio, el hecho de que a los sujetos de ambos grupos les hayan proyectado las mismas películas en inglés durante las clases, tiene una incidencia idéntica tanto en el grupo control como en el experimental, es decir, esta influencia extraña ha sido controlada y deja de ser una fuente de variación. Al influir por igual (+ 1) no añade ninguna diferencia o varianza, no afecta a la variación existente entre los datos que resultan de medir VD. Por tanto, si queremos ver la influencia de los MA V debemos controlar las otras fuentes de variación.

TABLA 5.3 Posibles fuentes de variación que pueden intervenir en el experimento de los MA V (Los datos son ficticios; sólo se exponen con fines didácticos para facilitar la comprensión de los conceptos; por tanto, se omite la significación estadística)

	(VD) Pretest	*(VI) MAV*	*Profesor*	*Memoria*	*Error medida*	*Películas*	*(VD) Postest*
GC	2	(+0)	(+1)	(+4)	(-1)	(+1)	7
	1	(+0)	(+1)	(-2)	(+2)	(+1)	3
	3	(+0)	(+1)	(+2)	(-2)	(+1)	5
	2	(+0)	(+1)	(-1)	(+3)	(+1)	6
	2	(+0)	(+1)	(+2)	(-2)	(+1)	4
	$\bar{X} = 2$						$\bar{X} = 5$
GE	1	(+3)	(+2)	(-1)	(+1)	(+1)	7
	4	(+3)	(+2)	(-2)	(-1)	(+1)	11
	1	(+3)	(+2)	(-1)	(+2)	(+1)	10
	2	(+3)	(+2)	(+1)	(-1)	(+1)	8
	2	(+3)	(+2)	(-1)	(+2)	(+1)	9
	$\bar{X} = 2$	Varianza primaria	Varianza secundaria	Varianza intragrupo o del error			$\bar{X} = 9$

↑ ↑ ↑ ↑ ↑ desviación sistemática de los datos ↓ ↓ ↓ ↓ ↓

Varianza sistemática | Varianza residual

El propósito principal del investigador es separar aquellas fuentes de variación en las que se halla interesado (VI) de las que no le interesan (variables extrañas). De ahí la necesidad de considerar dos fuentes de variación sistemática: la varianza primaria y la varianza secundaria.

- *Varianza primaria.* Desviación sistemática de los datos debida a la manipulación de VI. Cuanto más efectiva sea VI, más diferencia o desviación aparecerá entre la medición de VD efectuada en el grupo control y en el experimental. De ahí que se denomine *varianza intergrupos.*
- *Varianza secundaria.* Desviación sistemática de los datos debida a otras variables que no son VI y no han sido controladas. Por ello influyen de forma desigual, provocando varianza no pretendida.

B) *Varianza del error o aleatoria.* Constituida por el conjunto de fluctuaciones que presentan las medidas de la variable dependiente que tienden a actuar al azar. Posibles fuentes son:

- *Diferencias individuales.* Constituidas por características inherentes a los sujetos; por ejemplo, en el experimento de los MA V podría ser la memoria de cada sujeto. Ante un mismo tratamiento experimental no todos los sujetos reaccionan igual. Dado que las diferencias individuales tienden a desviar los datos en sentidos dispares, esta varianza se considera debida al error o aleatoria por ser impredecible.
- *Errores de medida.* Puede deberse a instrumentos poco precisos.
- *Procedimiento experimental.* La forma de aplicarlo puede ser diferente para los distintos sujetos.
- *Instrucciones.* Pueden darse de forma diferente por distintos experu:nentadores.

5.2 PRINCIPIO FUNDAMENTAL DE LA VARIANZA

En el experimento planteado sobre los MA V comprobaremos el siguiente principio fundamental de la varianza, es decir,

Varianza total = Varianza intergrupos + Varianza intragrupos

Como se recordará, los resultados medios de inglés, obtenidos en el postest, eran los siguientes:

Sin MAV			*Con MAV*
G.C.	Inglés	G	Inglés
a	7	f	7
b	3	9	**11**
e	5	h	10
d	6	i	8
e	4	j	9
	$\bar{X}=5$		$\bar{X}=9$

Sucesivamente procederemos a realizar el cómputo de la varianza total y de las varianzas intergrupo e intragrupo, con el [m de apreciar que aquélla se puede desglosar en la suma de estas últimas.

Cálculo de la varianza total. Variación de cada puntuación con respecto a la media total. Es decir:

$$S^2 = \frac{\Sigma (X - \bar{X})^2}{n} = \frac{60}{10} = \boxed{6}$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{70}{10} = 7$$

	X	$(X - \bar{X})$	$(X - \bar{X})^2$
	7	0	0
Grupo	3	−4	16
control	5	−2	4
	6	−1	1
	4	−3	9
	7	0	0
Grupo	11	4	16
experi-	10	3	9
mental	8	1	1
	9	2	4
Σ	70		60

Cálculo de la varianza intergrupos (primaria). La varianza total se ha calculado a partir de las diez puntuaciones en lengua inglesa; en cambio, la varianza intergrupo se refiere a la variación entre las medias de los grupos. En consecuencia, se utiliza la misma fórmula, pero en el cálculo sólo intervienen las medias 5 y 9 representando a cada grupo:

$$S^2 = \frac{\Sigma (X - \bar{X})^2}{n} = \frac{8}{2} = \boxed{4}$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{14}{2} = 7$$

	X	$(X - \bar{X})$	$(X - \bar{X})^2$
Grupo control	5	−2	4
Grupo experi.	9	+2	4
Σ	14		8

Cálculo de la varianza intragrupo (del error). Promedio de las variaciones calculadas dentro de cada grupo con respecto a las medias respectivas. El cómputo sería el siguiente:

$$S_c^2 = \frac{\Sigma (X - \bar{X})^2}{n} = \frac{10}{5} = \boxed{2}$$

$$\bar{X} = \frac{\Sigma x}{n} = \frac{25}{5} = 5$$

	X	$(X - \bar{X})$	$(X - \bar{X})^2$
	7	−2	4
Grupo	3	−2	4
control	5	0	0
	6	+1	1
	4	−1	1
Σ	25		10

De esta manera queda comprobado que se cumple el principio fundamental de la varianza:

$$S_e^2 = \frac{\Sigma(X-\bar{X})^2}{n} = \frac{10}{5} = 2$$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{45}{5} = 9$$

	X	$(X-\bar{X})$	$(X-\bar{X})^2$
Grupo experimental	7	−2	4
	11	+2	4
	10	+1	1
	8	−1	1
	9	0	0
Σ	45		10

$$S^2 = \frac{S_c^2 + S_e^2}{n} = \frac{2+2}{2} = \boxed{2}$$

De esta manera queda comprobado que se cumple el principio fundamental de la varianza:

6	=	4	+	2
Varianza total	=	Varianza intergrupos	+	Varianza intragrupos
Varianza total	=	Varianza sistemática	+	Varianza del error
Varianza total	=	Varianza primaria (VI) + Varianza secundaria (Extrañas)	+	Varianza residual (Dif. Individuales errores)

Las pruebas de significación, como el análisis de la varianza, se basan en el cociente:

$$F = \frac{\text{Varianza intergrupos}}{\text{Varianza intragrupos}} = \frac{4}{2} = 2$$

En la medida que aumente la varianza intergrupos (refleja el efecto de VI) y disminuya la varianza intragrupos (diferencias individuales y errores) el experimento será más eficaz y sensible, pues el cociente aumentará. Si nuestro experimento ficticio estuviera perfectamente controlado se verificaría que:

varianza secundaria = 0

pues las variables extrañas influirán por igual en cada grupo o se habrá eliminado su influencia. Por tanto, en el experimento ideal,

varianza intergrupos = varianza primaria = 4 (Sistemática) (VI)

Por tanto, en el buen experimento:

varianza intergrupo > varianza intragrupo

Para ello es necesario un control que consistirá en:

1) Maximizar la varianza primaria (producida por VI).
2) Neutralizar o eliminar la varianza secundaria (producida por las variables extrañas).
3) Minimizar la varianza del error (producida por las diferencias individuales y los errores de medida).

5.3 CONTROL DE LA V ARIANZA

El experimento es un tipo especial de investigación para examinar relaciones de causalidad. Para ello se introducen cambios deliberados en la variable independiente y se registran posibles modificaciones en la variable dependiente.

Con el fin de garantizar la validez del experimento, se utilizan técnicas de control de la varianza que pueden consistir en manipulación física, selectiva o estadística:

Manipulación física. Son técnicas de control directo que pueden ejercerse al aplicar la variable independiente o al igualar los efectos de algunas variables extrañas. Por ejemplo, espejos unidireccionales, suministrar drogas, premios, presentar a los sujetos estímulos verbales, luminosidad, ruidos o temperatura.

Manipulación selectiva. Algunos ejemplos pueden ser seleccionar y asignar aleatoriamente los sujetos, mantener constantes ciertas condiciones, aparear los sujetos, el bloqueo, el balanceo y el contrabalanceo.

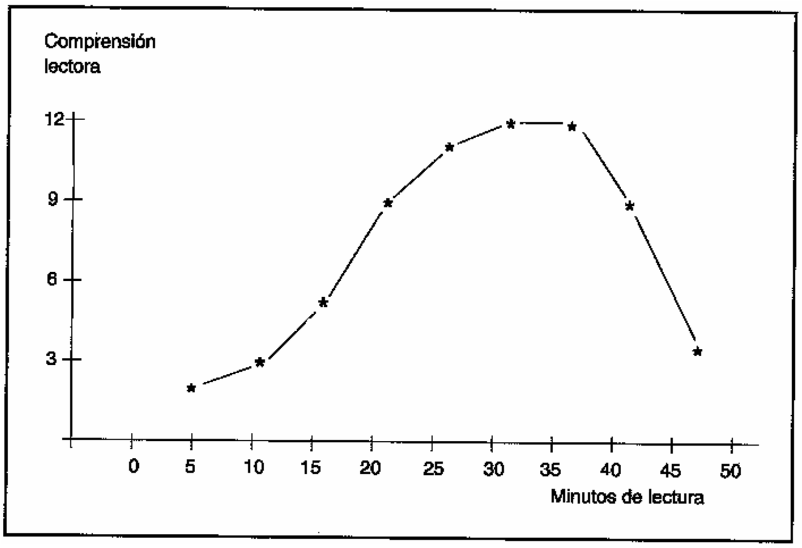
Manipulación estadística. Dada la naturaleza del fenómeno educativo es frecuente tener que recurrir a técnicas indirectas de control como son las técnicas estadísticas. Estas técnicas de control, como la correlación parcial y el análisis de covarianza, se utilizan cuando las variables no pueden ser sometidas a manipulación física o selectiva.

Como ya hemos apuntado más arriba, el investigador ha de recurrir a estas técnicas de control con el fin de maximizar la varianza primaria, eliminar la varianza secundaria

y minimizar la varianza del error. A continuación veremos más detenidamente distintas técnicas de control:

MAXIMIZACIÓN DE LA V ARIANZA PRIMARIA

Los niveles de la variable independiente han de adoptar en la realidad los valores decididos por el experimentador. Además, dichos valores han de ser extremos y óptimos. Supongamos que se realizan distintos experimentos en tomo a la eficacia de técnicas de estudio en niños de nueve años. Si queremos ver la influencia del tiempo de lectura en el rendimiento, habrá que asignar los valores adecuados. Así, valores como O, 1 Y 2 minutos de lectura pueden llevamos a concluir que el tiempo de lectura no influye, ya que no son valores extremos ni óptimos. Lo mismo podría ocurrir con valores más extremos, pero no óptimos, como O, 50 y 90 minutos, ya que al visualizar la relación entre dicha variable y una característica decisiva para el rendimiento (comprensión) se obtiene la figura 5.1.



FIG, 5.1. Distribución de la comprensión lectora según los minutos de lectura

En cambio, valores como O, 25 y 50 podrían ejercer un efecto más diferenciado, caso de existir una relación real. En consecuencia, podrían ser valores extremos y óptimos.

Existen tres formas de asignar los niveles de la variable independiente a los sujetos que permiten distintas modalidades de control: intergrupo, intragrupo y mixta.

Asignación intergrupo

Se utiliza esta modalidad cuando cada nivel de la variable independiente, también denominado *condición experimental* o *tratamiento,* se aplica a un grupo diferente de sujetos, es decir, cada sujeto es sometido a una sola condición experimental. Así, por ejemplo, para ver cómo influye el tiempo de estudio en la comprensión lectora de un texto T podemos seleccionar al azar 30 alumnos de B UP y asignarlos, también al azar, a un grupo control (GC) y dos grupos experimentales (GEl y GEl). Los alumnos asignados al GC no dedican tiempo a estudiar el texto, mientras que en cada GE se dedica 1 y 2 horas. Después de realizar el proceso descrito se comparan las medias de comprensión lectora referidas a dicho texto (tabla 5.4).

TABLA 5.4 Ejemplo de asignación de sujetos según la modalidad intergrupo

Grupo control (O horas)	*Grupo experimental* 1 (1 hora)	*Grupo experimental* 2 (2 horas)
1.Juan	11. Rosa	21. Lourdes
2. Ana	12. José	22. Carlos
3. Antonio	13. Raquel	23. Miguel
………….....	...	'''
9. Ramón	19. Mireia	29. Sergio
10. María	20. Víctor	30. Mónica
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$

En este experimento, comparar las medias tiene sentido si los tres grupos que intervienen y las tres situaciones ya eran homogéneas o equivalentes, antes de empezar el experimento y durante el desarrollo del mismo, especialmente en lo que se refiere a diferencias que pueden ser ocasionadas por variables extrañas que pueden influir en la variable dependiente. Para conseguir que estas variables afecten por igual a los tres grupos de sujetos y hacerlos comparables hay que controlar o igualar en cada grupo influencias como las siguientes:

- El experimentador: entusiasmo, amabilidad, etc.
- El contexto experimental: ruido, local, hora, instrucciones, etc.
- Los sujetos experimentales que presentan diferencias individuales en motivación, inteligencia, edad, sexo, capacidad de estudio, etc.

Se consigue un control ideal cuando lo único que diferencia a cada grupo es el tratamiento proporcionado a través de los niveles de la variable independiente (O horas, 1 hora y 2 horas). En consecuencia, si antes de empezar el experimento los tres grupos presentaban medias estadísticamente iguales en la comprensión lectora del texto T,

puede concluirse, con más seguridad, que si aparece alguna diferencia entre las medias después del experimento, dicha diferencia podrá ser atribuida a que los grupos han sido sometidos a distintos niveles de la variable independiente.

Asignación intragrupo

Según esta modalidad todos los niveles de la variable independiente se aplican sucesivamente al mismo grupo de sujetos, es decir, cada sujeto es sometido a todas las condiciones experimentales. En el ejemplo anterior, podemos someter a los 30 sujetos a la prueba de comprensión lectora sobre el texto T para obtener la media cuando el nivel de la variable independiente vale O horas de estudio. En dos días sucesivos el grupo total estudia durante 1 hora y 2 horas respectivamente, midiéndose la comprensión después de cada período de estudio (tabla 5.5).

TABLA 5.5 Ejemplo de asignación de sujetos según la modalidad intragrupo

Grupo total (0 horas)	*Grupo total* (1 hora)	*Grupo total* (2 horas)
1. Juan	1. Juan	1. Juan
2. Ana	2. Ana	2. Ana
3. Antonio	3. Antonio	3. Antonio
…	…	…
...		...
29. Sergio	29. Sergio	29. Sergio
30. Mónica	30. Mónica	30. Mónica
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}$

En este caso no hay tanta necesidad de controlar variables extrañas que proceden de los sujetos. Al ser los mismos, dichas variables afectarán por igual a los resultados en comprensión lectora. En cambio, hay que controlar las variables vinculadas al experimentador y , especialmente, las que están implicadas en el contexto experimental y son inherentes a esta modalidad de asignación de los tratamientos. Así, hay que tener en cuenta que:

1) Los efectos de los niveles de la variable independiente (tiempo de estudio) pueden permanecer desde una aplicación hasta las siguientes acumulándose. En este sentido, para controlar este efecto no pretendido podría recurrirse a tres textos de dificultad similar pero con diferente contenido.

2) Que los sujetos pueden mejorar su comprensión lectora por el hecho de pasarles tres pruebas sobre comprensión.

3) Los niveles de VI pueden tener diferente efecto según el orden de aplicación.

Aunque existen técnicas de control que tratan de igualar la influencia de estas variables extrañas, para el experimento planteado es más viable una aplicación intergrupo; sin embargo, el mismo ejemplo ha sido incluido aquí con fines meramente expositivos.

Asignación mixta

En esta modalidad a varios sujetos se les aplica solamente un nivel de la variable independiente (intergrupos) y a otros sujetos se les aplican varios niveles de la variable independiente (intragrupo). Así, en el ejemplo mencionado, después de aplicar la prueba de comprensión al grupo total, el mismo grupo de 30 sujetos puede estudiar durante 1 hora, mientras que sólo 10 sujetos estudian durante 2 horas (tabla 5.6).

TABLA 5.6. Ejemplo de asignación de sujetos según la modalidad mixta

Grupo total (O horas)	*Grupo total* (1 hora)	*Grupo total* (2 horas)
1. Juan	1. Juan	1. Rosa
2. Ana	2. Ana	2. José
3. Antonio	3. Antonio	3. Raquel
...	...	...
...		
29. Sergio	29. Sergio	29. Mireia
30. Mónica	30. Mónica	30. Víctor
$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$

Esta manera de asignar los tratamientos combina las ventajas de las condiciones intergrupo e intragrupo.

ELIMINACIÓN DE LA V ARIANZA SECUNDARIA

El experimentador ha de garantizar que los cambios observados en la variable dependiente (varianza) sean únicamente debidos a los distintos niveles de la variable independiente. Para ello las variables extrañas no han de contaminar los resultados finales del experimento, es decir, las variables extrañas han de incidir con la misma intensidad en los grupos experimentales y en el grupo control. De esta forma, los grupos experimentales y el grupo control sólo se diferenciarían en que reciben distintos niveles de la variable independiente. En consecuencia, las diferencias observadas en la variable dependiente serán atribuibles a los niveles de la variable independiente siempre que las variables extrañas no influyan de forma diferenciada.

Si el experimento no se controla adecuadamente, existe la posibilidad de que los

cambios observados en la variable dependiente (varianza total) puedan ser provocados, en parte, por una gran multiplicidad de variables extrañas que junto a la variable independiente pueden producir modificaciones en la variable dependiente. Las variables extrañas que pueden afectar a los resultados finales han sido englobadas en categorías como las siguientes:

- Sujetos experimentales: edad, sexo, capacidad cognitiva, personalidad, destreza, etc.
- Contexto experimental: instrucciones, aparatos, ruidos, etc.
- Experimentador: expectativas, actitudes, *rol,* etc.

De ahí que el investigador deba cerciorarse de que las variaciones que aparecen en la variable dependiente puedan ser atribuidas exclusivamente a las variaciones introducidas deliberadamente a través de la variable independiente. Para ello puede utilizar distintas técnicas de control.

Las principales formas y técnicas de *control de las variables extrañas* se basan fundamentalmente en la *aleatorización* y la selección. La *aleatorización* constituye el procedimiento más adecuado para controlar las variables extrañas conocidas y desconocidas que proceden de los sujetos. Gracias a la *selección* un mismo método didáctico puede utilizarse para explicar a los alumnos diferentes conceptos que pueden seleccionarse según su dificultad o adecuación a la edad. Es posible formar grupos según el sexo, edad, clase social, motivación, inteligencia o personalidad cuando sospechamos o sabemos que estas variables pueden tener una influencia significativa. Mediante la selección se pueden eliminar variables, como pueden ser el ruido, o bien elegir un valor único o una única categoría: niños o niñas, período evolutivo, procedencia, etc., teniendo en cuenta que limitaremos el poder generalizador de las conclusiones. Algunas variables contextuales, como la temperatura y la hora de aplicación de los instrumentos de medida, también podemos mantenerlas constantes.

A continuación se detalla la técnica de la aleatorización y diversos procedimientos, algunos de ellos basados en la selección:

Aleatorización

Consiste en asignar los sujetos al azar a cada grupo del experimento, lo que permite igualar todas las diferencias individuales (carácter, inteligencia, motivación, comprensión lectora, etc.). La asignación aleatoria asegura que dichas variables tengan la misma oportunidad de influir en la variable dependiente, llegando a compensarse sus efectos y resultando grupos equivalentes antes de aplicarse la variable independiente.

La aplicación estricta de esta técnica requiere las siguientes fases (fig. 5.2):

1) Seleccionar al azar una muestra de la población.

2) Asignar los sujetos al azar a tantos grupos como niveles tenga la variable independiente. Por ejemplo, grupo 1 (G_1) y grupo 2 (G_2).

3) Formados los grupos del experimento se asigna al azar cada nivel de la variable independiente a un grupo. El grupo que reciba el nivel mínimo de la variable independiente o ausencia de tratamiento experimental (T_o), se denominará grupo control

(GC), Y grupo experimental (GE) el que reciba el tratamiento (TI). De ahí que estas dos últimas fases sean más propias de la modalidad de asignación intersujetos.

Cuando el número de sujetos de la muestra es suficientemente grande, la probabilidad de que las diferencias individuales se compensen unas a otras, por efecto del azar, es también mayor, existiendo más garantía de que los grupos sean equivalentes antes de iniciarse el experimento. Como es evidente, aquellas variables extrañas que puedan influir o generarse después de haberse formado los grupos (cansancio, ansiedad, experimentador, ambiente, contexto experimental, etc.) no pueden controlarse mediante la técnica de la aleatorización.

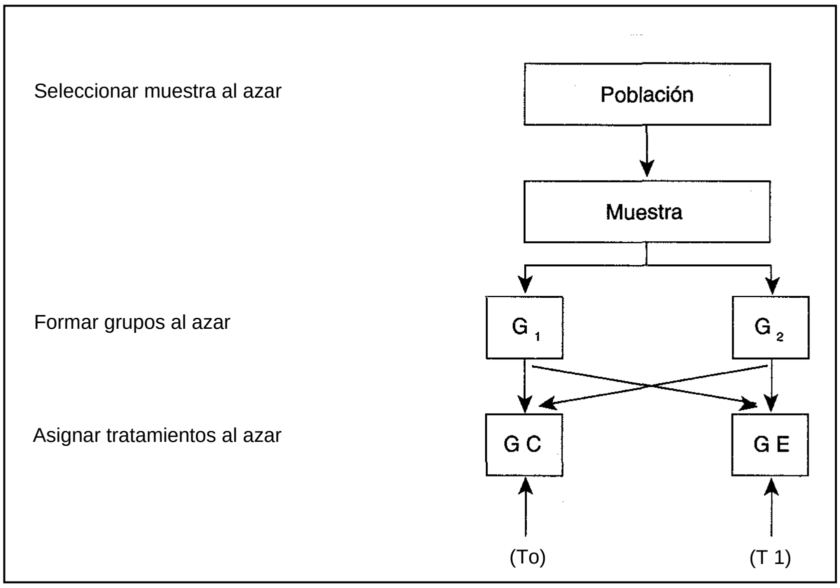


FIG. 5. 2 Fases de la aleatorización como técnica de control (adaptado de Rosel, 1986,244)

Constancia

Las variables extrañas se mantienen a un nivel constante. Por ejemplo, si queremos ver la eficacia de distintos métodos didácticos se procurará que para todos los grupos del experimento los temas que hay que explicar sean los mismos. También convendrá que otras características, como el nivel de iluminación del local, la claridad al especificar las instrucciones, etc., intervengan por igual en cada grupo del experimento. La constancia (y también la eliminación) presenta el inconveniente de que restringe la generalización de las conclusiones del experimento. Así, cuando la edad se mantiene constante a los 13 años o cuando sólo se seleccionan chicos, los resultados del experimento no serán extensibles a los alumnos de 10 años ni a las chicas, respectivamente.

Eliminación

Es un caso especial de constancia que se lleva a cabo cuando algunas condiciones ambientales se restringen al máximo, manteniéndolas constantes pero a nivel de ausencia, es decir, a nivel cero. Por ejemplo, pueden eliminarse elementos distorsionantes como el ruido, la visión de objetos que distraen, etc., lo que hace que ignoremos qué ocurriría si el experimento se llevara a cabo en condiciones normales.

Apareo

Cuando sospechamos que una variable extraña (variable de apareo) está correlacionada y/o puede influir decisivamente en la variable dependiente, podemos formar grupos equivalentes antes de aplicar la variable independiente, recurriendo al apareo de los sujetos. Para ello asignamos a cada grupo del experimento un sujeto que posea la misma magnitud en la variable de apareo o de bloqueo.

Supongamos que queremos comprobar la eficacia de un programa de informática para resolver ecuaciones. Podemos utilizar la inteligencia como variable de apareo, después de comprobar que se halla fuertemente correlacionada con la habilidad para resolver ecuaciones. En la práctica, aunque se utilice la aleatorización, cuando la muestra no es muy grande, puede darse el caso de que el grupo experimental esté constituido por sujetos más inteligentes. Si así ocurriera, el posible aumento en la habilidad para resolver las ecuaciones podría deberse no sólo al programa de informática, sino también a la inteligencia. Así, para aplicar la técnica del apareo en el ejemplo propuesto podríamos aplicar el test de inteligencia «D-48», lo que nos permitiría formar un grupo control (GC) y otro experimental (GE), agrupando por parejas los sujetos más afines en dicha variable y asignándolos al azar a cada grupo, lo que origina distintos pares o bloques (tablas 5.7 y 5.8).

Como los sujetos dentro de cada bloque o par tienen la misma inteligencia, resultará que los grupos control y experimental serán equivalentes antes de aplicar el programa de informática. Como es evidente, la propia variable dependiente también podría utilizarse como variable de apareo.

TABLA 5. 7 Puntuaciones de inteligencia «D-48»

1. Eva	30	7. Ignacio	35
2. Pablo	33	8. Rosa	28
3. Carlos	34	9. Javier	33
4. Gemma	35	10. Ana	30
5. Manuel	28	11. Marco	27
6. Carmen	27	12. Sonia	34

TABLA 5.8 Pares de sujetos y asignación a cada grupo

Pares	GC	*GE*
1.	Ignacio	Gemma
2.	Carlos	Sonia
3.	Pablo	Javier
4.	Eva	Ana
5.	Rosa	Manuel
6.	Carmen	Marco
	X_0	X_j

En la representación simbólica de un diseño basado en la técnica del apareo los distintos pares pueden representarse por:

Ap = asignación según apareo

En general, esta técnica de control presenta inconvenientes como los siguientes:

- Cuando la variable de apareamiento presenta gran variabilidad hay que tener más flexibilidad en el grado de semejanza que han de tener las puntuaciones para incluidas en el mismo par.
- Con frecuencia, hay que prescindir de algunos sujetos si no tienen un «par» adecuado que presente un nivel similar en la variable de apareamiento.
- De la misma manera, cuando ya se han formado los pares y falla un sujeto por enfermedad o absentismo, hay que eliminar el par si no existe otro sujeto similar que pueda sustituir al ausente.
- Si se quiere aparear a los sujetos a partir de más de una variable de apareo, las posibilidades de formar los pares son menores y la muestra quedaría muy reducida.

Bloqueo

Cuando se forman más de dos grupos de tratamiento, en lugar de agrupar por pares interesa formar bloques a partir de una o más variables de bloqueo. Esta técnica presenta inconvenientes similares a los que se han expuesto al considerar la técnica del apareo y hay que admitir un mayor margen de variabilidad en los sujetos que se agruparán en un mismo bloque; de lo contrario habría que prescindir de muchos sujetos. Sin embargo, como dentro de cada bloque los sujetos se asignan también al azar a cada grupo de tratamiento, es fácil conseguir grupos equivalentes con respecto a la variable de bloqueo.

En el ejemplo planteado en el apartado anterior, si queremos comprobar la eficacia relativa de tres programas de informática para resolver las ecuaciones, podemos formar un grupo control y tres grupos experimentales, utilizando las puntuaciones de inteligencia para agrupar tres bloques de inteligencia alta, media y baja (tabla 5.9).

Balanceo

Esta técnica distribuye equitativamente la presencia de una variable extraña en los grupos que deseamos formar. Si en el experimento mencionado anteriormente sobre la eficacia de los programas de informática se sospecha que las chicas pueden tener una mayor habilidad para la resolución de ecuaciones, no sabríamos si la posible mejora observada en los grupos experimentales es debida a los programas, al hecho de que en dichos grupos pudiera predominar el sexo femenino o a ambas circunstancias.

TABLA 5.9 Ejemplo ilustrativo de la técnica de bloqueo

Bloques autorizados	*GC*	*GE₁*	*GE₂*	*GE₃*
1. Inteligencia alta	Ignacio	Carlos	Gemma	Sonia
2. Inteligencia media	Eva	Javier	Ana	Pablo
3. Inteligencia baja	Carmen	Marco	Rosa	Manuel
	$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$

Supongamos que al comparar las medias se obtiene que bajo el programa 2 (GE2) la habilidad para resolver las ecuaciones es superior a la que presentan los sujetos que han utilizado el programa 1 (GE1). Esta diferencia entre las medias obtenidas podría deberse a que el programa 2 es más eficaz que el programa 1 y/o al hecho de que el grupo GE2 está constituido por chicas y el grupo GE1 por chicos. En consecuencia, los resultados del experimento quedarían contaminados por la variable sexo. Hay que tener en cuenta que la muestra debería ser más grande y entonces sería menos probable que predominara uno de los sexos en algún grupo después de aleatorizar la asignación dentro de cada bloque. Una posible solución sería formar los grupos incluyendo la misma cantidad de chicos y chicas en cada grupo. Para recoger la mayor información posible con respecto a los efectos de la variable sexo, y suponiendo que no interese controlar la inteligencia, el experimento podría diseñarse según la técnica del balanceo para controlar la variable sexo. Para una mayor sencillez veamos cómo podría comprobarse la eficacia de un solo programa (tabla 5.10).

TABLA 5. 10 Ejemplo con la utilización de la técnica del balanceo

		GC	GE	
Sexo	M	Ignacio Carlos Pablo	Javier Marco Manuel	$\bar{X}_M$
	F	Eva Carmen Gemma	Ana Rosa Sonia	$\bar{X}_F$
		$\bar{X}_0$	$\bar{X}_1$	$\bar{X}_T$

Como puede observarse, el experimento tiene cuatro subgrupos de sujetos: dos subgrupos forman el grupo control (tres chicos y tres chicas) y el programa es aplicado a otros dos subgrupos (tres chicos y tres chisas) y dos subgrupos. Si se incluyen más sujetos en cada subgrupo, la comparación $\bar{X}_0$ y $\bar{X}_1$ nos informa sobre el efecto del programa, ya que en el grupo control no ha sido aplicado. Además, podemos obtener una información adicional sobre el sexo al comparar las medias $\bar{X}_M$ y $\bar{X}_F$. En realidad, hemos sistematizado o equilibrado una variable extraña (sexo), considerándola como otra variable independiente de la investigación.

Contrabalanceo o equiponderación

Con esta técnica, los grupos se dividen en pequeños sub grupos con el fin de equilibrar los efectos de las variables extrañas. Puede utilizarse en la modalidad de asignación intragrupo, es decir, cuando distintos tratamientos aplicados sobre los mismos sujetos pueden producir efectos distintos según la secuencia u orden de aplicación. Supongamos que hemos elaborado varias reglas nemotécnicas para potenciar las facultades y alcance de la memoria de un grupo de niños débiles mentales y pretendemos comprobar la eficacia relativa de dos técnicas (T) y (T) para memorizar frases. Como es obvio, al aplicar cada técnica (VI) y al contabilizar cuántas frases han memorizado (VD), se recurriría a frases distintas pero de dificultad similar. Sin embargo, en cada aplicación habría que tener en cuenta dos variables extrañas que pueden generarse durante el experimento:

1) Cierto nivel de aprendizaje práctico, experiencia o habilidad que puede facilitar la memorización de las frases después de aplicar la primera técnica.

2) La fatiga, que puede favorecer que el sujeto cometa más errores a medida que aumenta el número de tratamientos.

El grado de memorización de las frases dependerá no sólo de la eficacia de la técnica aplicada, sino también de la acción conjunta del aprendizaje y de la fatiga, que originan el denominado *error progresivo.* Aunque éste puede disminuir con adecuados intervalos intertratamientos, ya que el tiempo suele eliminar los efectos de dicho error progresivo, es más viable neutralizarlo por medio del contrabalanceo (tabla 5.11).

TABLA 5. 11 Técnica de contrabalanceo. Modalidad intragrupo

T_1 (1.ª)	T_2 (2.ª)	T_2 (3.ª)	T_1 (4.ª)
Alicia	Alicia	Javier	Javier
Iván	Iván	Isabel	Isabel
Roberto	Roberto	Laura	Laura
...	...	...	...
...	...	...	...
Olga	Olga	Irene	Irene
Ramón	Ramón	Vicente	Vicente
$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$

La equiponderación o contrabalanceo puede adoptar también la modalidad intergrupo (tabla 5. 12).

TABLA 5. 12 Técnica de contrabalanceo. Modalidad intergrupo

T_1 (1.ª)	T_2 (2.ª)	T_2 (3.ª)	T_1 (4.ª)
Alicia	Alicia	Javier	Javier
Iván	Iván	Isabel	Isabel
Roberto	Roberto	Laura	Laura
...	...	...	...
...	...	...	...
Olga	Olga	Irene	Irene
Ramón	Ramón	Vicente	Vicente
$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$

En ambas modalidades se ha incluido entre paréntesis la secuencia de aplicaciones y la media del número de frases memorizadas después de haber practicado con cada técnica. Así, los niveles medios que permitirán estimar la eficacia relativa de cada regla nemotécnica podrían ser:

$$X_{T1} = \frac{\bar{X}_1 + \bar{X}_2}{2} \qquad X_{T2} = \frac{\bar{X}_3 + \bar{X}_4}{2}$$

ANÁLISIS DE COV ARIANZ

Las técnicas de control descritas pueden plantear al educador dificultades como las siguientes:

1) Formar nuevos grupos para el experimento, por ejemplo, a través de la aleatorización, puede ser no recomendable por razones organizativas de los centros. Esto obliga a utilizar grupos naturales, respetando las agrupaciones que constituyen las distintas clases o conglomerados.
2) La utilización de otras técnicas como el apareo y los bloques puede resultar inapropiada, ya que el educador suele disponer de pocos sujetos para formar grupos homogéneos igualados en variables extrañas relevantes.

En consecuencia, al tomar grupos ya constituidos (naturales), las variables extrañas pueden estar presentes con distintos niveles de intensidad, originando diferencias entre los grupos antes de empezar el experimento. Si estas diferencias iniciales no se pudieran controlar (igualar) mediante algún sistema alternativo de control, la investigación no podría realizarse. Por ejemplo, podría darse la situación ficticia de la tabla

5.13 al aplicar dos métodos de enseñanza (M_1 y M_2) a dos grupos de tercero de primaria y analizar sus efectos sobre el rendimiento tomando un grupo control (GC).

TABLA 5.13 Datos de una investigación en la que no se ha controlado la inteligencia

Grupos naturales	*Número sujetos*	*Inteligencia media (0-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
3.ºA	26	30	2	GC	4
3.º B	29	28	1	M_1	7
3.º C	27	32	3	M_2	8

Supongamos que la inteligencia está significativamente correlacionada con el rendimiento, y no podemos controlarla mediante las técnicas mencionadas. El análisis de covarianza permite un control estadístico o control indirecto de esta variable extraña al tener en cuenta las diferencias iniciales entre tercero A, tercero B y tercero C. Gracias al análisis de covarianza el experimento puede realizarse. Esta técnica estadística ajusta las diferencias finales con respecto a las iniciales y compara las medias tomando en consideración la posible existencia de diferencias iniciales que podrían enmascararlas. Es decir, permite corregir o «ajustar» los cambios observados en la variable dependiente, dejando solamente los efectos atribuibles a la variable independiente.

La situación planteada es comparable a lo que ocurriría con tres niños en un concurso para coleccionar cromos. Si el educador quiere saber quién es más hábil coleccionando y sospecha que los niños empiezan el concurso con distinto número de cromos, habrá de comparar las cantidades finales acumuladas tomando en consideración la cantidad inicial que ya tenía cada niño. Es decir, ha de «ajustar» o hacer comparables las cantidades finales. La pregunta implícita sería: ¿cuántos cromos habría ganado cada niño si hubieran empezado sin ninguno o con la misma cantidad?

En el caso que nos ocupa el análisis de covarianza resuelve estadísticamente este interrogante y compara las medias que tendría después cada grupo si inicialmente las hubieran tenido iguales. Por ejemplo, las medias definitivas comparables o ajustadas hubieran podido ser 4, 8 y 7, lo que podría sugerir cierta supremacía del método M 1 (tabla 5.14):

TABLA 5. 14 Datos ficticios de una investigación en la que se ha controlado la inteligencia ajustando las medias a través de un análisis de covarianza

Grupos naturales	*Número sujetos*	*Inteligencia media (0-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
6.º A EGB	26	30	2	GC	4
6.º B EGB	29	28	1	M_1	8
6.º C EGB	27	32	3	M_2	7

En cierto modo, la aportación del análisis de covarianza sería similar a la información que ofrecería el siguiente diseño si en el caso planteado se hubiera podido controlar la inteligencia aleatorizando o seleccionando grupos de sujetos equivalentes (tabla 5.15).

TABLA 5.15 Datos de una investigación en la que se ha controlado la inteligencia directamente aleatorizando o seleccionando grupos equivalentes

Grupos naturales	*Número sujetos*	*Inteligencia medía (D-48)*	*Rendimiento medio antes*	*Aplicación métodos*	*Rendimiento medio después*
3.º A EGB	26	30	2	GC	4
3.º B EGB	29	30	2	M,	8
3.º C EGB	27	30	2	M_2	7

La variable extraña (inteligencia) que se distribuye de forma desigual en los grupos del experimento recibe el nombre de *covariable.* El análisis de covarianza se basa en el supuesto de que el nivel en rendimiento obtenido por cada sujeto depende del valor que tenga en inteligencia y del método que se le aplique. En consecuencia, la media en rendimiento se reajusta en función de los métodos de enseñanza y en función de la covariable inteligencia.

MINIMIZACIÓN DE LA V ARIANZA INTRAGRUPO

Como se recordará, la varianza del error, residual o intragrupo se refería a variaciones debidas al azar. También habíamos contemplado algunas fuentes de esta varianza, como son las diferencias individuales y los errores de medida debidos a la imprecisión de los instrumentos utilizados en el proceso de medición.

Para minimizar la varianza del error conviene utilizar instrumentos fiables al medir la variable dependiente, reducir las diferencias individuales, aplicar bien los niveles de la variable independiente y comunicar las mismas instrucciones a los sujetos que intervienen en el experimento.

5.4 FUENTES DE INVALIDEZ

Hemos visto que en las investigaciones hay que tener en cuenta tres tipos de validez: la interna, que exige el control; la externa (generalización) y la conceptual (coherencia). En el caso de la investigación experimental estos tipos de validez revisten una importancia decisiva. Las variables o elementos que ponen en peligro la validez del experimento se denominan *fuentes de invalidez.* Por su relevancia nos centraremos en las fuentes de invalidez interna y externa.

FUENTES DE INVALIDEZ INTERNA

Cuando un experimento carece de validez interna, el investigador no tiene garantías de que las diferencias observadas (varianza total) en la variable dependiente puedan atribuirse exclusivamente a la aplicación intencional de los distintos niveles de la variable independiente. Existe la posibilidad de que dicha variación sea debida a la influencia de otras variables extrañas o fuentes de invalidez, es decir, hay motivos para sospechar que las variables extrañas no influyen por igual en todos los grupos del experimento, ya que no se aplican técnicas de control adecuadas. En consecuencia, cualquier elemento que afecte al grado de control del diseño de investigación constituye una fuente potencial de invalidez interna.

Las variables extrañas, que de no controlarse pueden constituir fuentes de invalidez interna, las agruparemos bajo las siguientes denominaciones genéricas:

a) *Historia.* Engloba cualquier acontecimiento externo a los sujetos y diferente de la variable independiente que influya a lo largo del experimento (entre el pretest y el postest) en la variable dependiente. Estos elementos distorsionantes pueden producirse dentro o fuera del marco del experimento. El efecto puede agudizarse cuando

1) el tratamiento es largo;
2) se trata de un experimento de campo;
3) los grupos no reciben los tratamientos simultáneamente.

b) *Maduración.* Consiste en un cambio que se produce en los sujetos debido a su desarrollo biológico o psicológico al transcurrir el tiempo. Puede estimarse a través del cambio observado en un grupo control.

c) *Selección diferencial de los sujetos.* Puede ocurrir en investigaciones cuasiexperimentales, cuando no se forman al azar los distintos grupos que intervienen en el diseño. Así, por ejemplo, cuando el investigador ha de utilizar grupos ya constituidos como aulas de un colegio. Dado que en cada grupo las variables extrañas han podido influir previamente de forma diferenciada, no puede garantizarse que los grupos sean homogéneos antes de aplicar la variable independiente. En consecuencia, los cambios observados pueden ser debidos a las diferencias existentes entre los grupos más que a la acción del tratamiento.

d) *Pérdida de sujetos.* Esta fuente de invalidez, también denominada *mortalidad experimental,* puede afectar especialmente si la pérdida se produce en los sujetos situados en los extremos de la distribución.

e) *Interacciones.* Pueden aparecer entre selección y maduración, selección e historia, etc. El grupo experimental puede hallarse más motivado, o ser más ágil al entender las instrucciones, y por ello, independientemente del tratamiento, puede obtener mayor puntuación que el grupo control en el postest.

f) *Reactividad de la medida.* Las medidas o la aplicación de pruebas puede afectar a los sujetos, o pueden responder éstos de forma diferente a como lo harían en condiciones normales, lo que influiría en los resultados finales, independientemente de la acción de la variable independiente.

g) *Instrumentación.* Ciertos instrumentos de medida pueden cambiar con el tiempo. Por ejemplo, algunos sujetos pueden hacer marcas en los protocolos de

algunas pruebas que condicionarán las respuestas de otros sujetos si se utilizan los mismos protocolos, o un observador puede cambiar de criterio durante la recogida de datos o durante el análisis de los mismos. También puede ocurrir que los instrumentos de medida no tengan una fiabilidad adecuada.

h) *Regresión estadística.* Cuando se aplica una prueba en sucesivas ocasiones (pretest-postest) las puntuaciones progresivas tienden a aproximarse a la media del grupo de procedencia. Este efecto es más pronunciado cuando se seleccionan grupos en base a puntuaciones extremas. Por ejemplo, si queremos probar la eficacia de un programa de técnicas de estudio para los mejores alumnos de un curso de primaria podemos aplicar un test de rendimiento con el fin de seleccionar los alumnos que presentan las puntuaciones más altas. Si se les pasa posteriormente un nuevo test de rendimiento, independientemente de la aplicación de las técnicas de estudio, la media que obtendrá el grupo en este segundo test tenderá a desplazarse hacia la media de la población original, es decir, tiende a descender. Si se hubieran seleccionado los alumnos más retrasados, la segunda vez que se pasara el test la media tendería a incrementarse. En ambos casos ocurre un fenómeno de regresión estadística, principalmente debido al azar y a posibles deficiencias de los instrumentos.

La regresión estadística produciría el efecto que aparece en la figura 5.3. Supongamos que al aplicar el primer test el grupo total obtiene una media global de 5, el grupo de alumnos más retrasado sólo registra una media de 3 y los más adelantados de 8. Si realizamos la segunda aplicación de un test de rendimiento, cubriendo el mismo contenido y sin introducir ningún cambio en los sujetos, la media del grupo más retrasado tenderá a aumentar, aproximándose a la media global del grupo de procedencia (5). Por el contrario, la media del grupo más adelantado tenderá a descender aproximándose a la misma media global. Así, si un grupo con puntuaciones altas en alguna variable se utiliza como grupo control, y otro con puntuaciones bajas se constituye en experimental, la regresión no está controlada y puede actuar favoreciendo a los sujetos experimentales.

i) *Difusión del tratamiento.* Los sujetos del grupo experimental pueden comunicar a los demás las características del tratamiento.

FUENTES DE INVALIDEZ EXTERNA

Las fuentes de invalidez externa son condiciones que afectan a la representatividad o generalización de los resultados de una investigación, es decir, a la posibilidad de generalizar los resultados a otros sujetos, a otros grupos y a otras condiciones y situaciones reales.

Las fuentes de invalidez externa más importantes que pueden afectar a la representatividad de los resultados son:

a) *Reactividad experimental.* Cuando los sujetos sometidos a las condiciones experimentales actúan de diferente forma de como lo hacen en condiciones normales. Esta variable puede afectar también a la validez interna.

b) *Interacciones.* Pueden producirse entre los errores de selección y la variable experimental. Cuando se realiza un experimento sobre una muestra sesgada (voluntarios, sujetos de un determinado nivel cultural, social, etc.), los resultados no podrán

generalizarse a otros grupos diferentes no representados en la muestra. De la misma manera, si el experimento se lleva a cabo en un contexto muy específico (laboratorio, colegio, centro de rehabilitación, etc.), no existen suficientes garantías de que los resultados puedan extrapolarse a otros ámbitos. En general, para que los resultados puedan generalizarse a situaciones de la vida real, el experimento ha de realizarse en condiciones similares a contextos reales.

c) *Efecto reactivo o interactivo del pretest.* La aplicación de un test previo puede restringir el alcance de la generalización de los resultados obtenidos. El pretest puede entrenar a los sujetos en relación con la variable independiente; así, los resultados obtenidos no serán generalizables a un grupo de sujetos que no hayan pasado el pretest.

d) *Inteiferendas de tratamientos múltiples.* Cuando se administra una serie de tratamientos, a menos que el efecto de los tratamientos anteriores quede eliminado con el tiempo, no se sabrá si los sujetos reflejan el efecto del tratamiento que se les aplica en ese momento o la posible interacción existente entre dicho tratamiento y los efectos de los tratamientos anteriores.

e) *Novedad del tratamiento.* La curiosidad, motivación o novedad pueden influir «aumentando» el efecto de la variable independiente. Cuando se aplicara el mismo tratamiento en un contexto real perdería efectividad a medida que se prolongara el tratamiento.

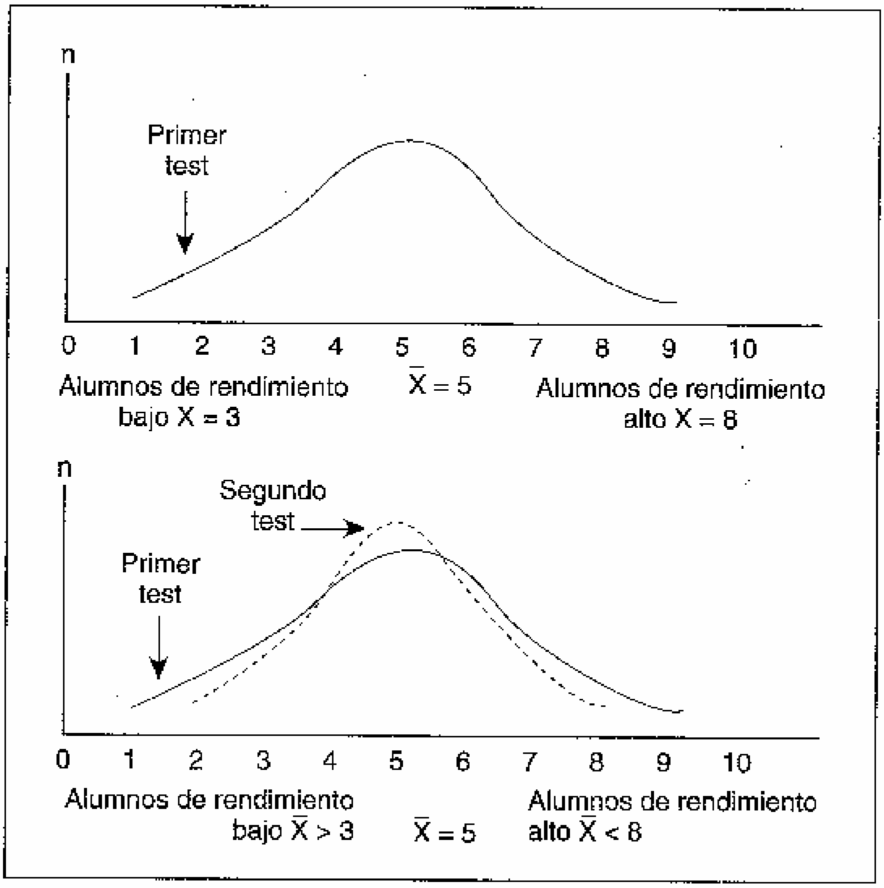


FIG. 5. 3 Efecto de la regresión estadística cuando se aplica un test por segunda vez en grupos extremos

5.5 CRITERIOS DE CLASIFICACIÓN DE LOS DISEÑOS

El método experimental puede adoptar diversas modalidades, siendo muchos los criterios de clasificación que pueden utilizarse: técnicas de control (Amau, 1978), número de grupos (Jiménez, 1985), mayor o menor rigurosidad, validez interna y externa (Campbell y Stanley, 1979), número de variables de tratamiento (Kirk, 1982). La adopción de uno u otro tipo de criterio genera taxonomías diferenciadas.

Didácticamente, el criterio del control parece ser el que inicialmente ofrece más ventajas, al desglosar los diseños en tres categorías: grupos al azar, bloques homogéneos e intragrupo. Dentro de cada categoría pueden efectuarse subdivisiones según el criterio del número de variables independientes y del número de niveles de tratamiento; así ocurre con los experimentos bivalentes, multivalentes y factoriales.

En la figura 5.4 pueden verse los diseños más importantes con asignación intergrupo o intragrupo y clasificados según el número de variables y las técnicas de control.

Para elegir el diseño que mejor se adapte a las necesidades del investigador y del problema planteado hay que tener en cuenta criterios como los siguientes:

1) Número de variables independientes que intervienen en la investigación. Cuando interviene una sola variable independiente hay que optar por los diseños unifactoriales o simples, y si están implicadas más de una variable independiente hay que elegir entre los multifactoriales o simplemente denominados factoriales.

2) Número de variables dependientes. Nuestra exposición se centrará siempre en diseños que contemplen una sola variable dependiente.

3) La asignación de los sujetos a los niveles de la variable independiente. Podemos elegir entre la asignación intergrupo e intragrupo. Si se opta por la asignación intergrupo podemos formar los grupos completamente al azar o según bloques. Como puede apreciarse en la figura 5.4, las diferencias individuales serán máximas en los diseños completamente al azar y mínimas en los diseños intragrupo.

4) Número de niveles de la variable independiente. Cuando dicha variable tiene dos niveles se utiliza un diseño bivalente, y si tiene más de dos se utiliza un diseño multivalente.

En los apartados siguientes se describen brevemente los diseños más utilizados.

5.6 DISEÑOS COMPLETAMENTE AL AZAR

Después de elegir la muestra al azar, en estos diseños los grupos se forman asignando al azar cada sujeto a uno de los grupos del experimento, y posteriormente, asignando también aleatoriamente cada nivel de tratamiento a los distintos grupos. Para que el azar pueda garantizar la equivalencia de los grupos hay que recurrir a muestras suficientemente grandes. De esta manera puede esperarse que las posibles variables contaminadoras se compensen y afecten por igual a todos los grupos antes de empezar los tratamientos.

Dentro de esta modalidad los diseños más utilizados son los bivalentes, multivalentes y factoriales.

Disminuyen las diferencias individuales →
Se reduce el error experimental

CRITERIOS / DISEÑOS	1.º Núm. VI	2.º Núm. VD	3.º ASIGNACIÓN EXPERIMENTAL: INTERGRUPOS – Completamente al azar	INTERGRUPOS – De bloques	INTRAGRUPOS
UNIFACTORIAL O SIMPLE	1	1	* Bivalente: Pretest - postest, Sólo postest * Multivalente: Pretest- postest, Sólo postest	* Grupos apareados * Bloqueo simple	* Tratamientos por sujetos * Análisis de tendecias
MULTIFACTORIAL O FACTORIAL	> 1	1	* Factorial	* Factorial bloques al azar	* Factorial intragrupo
* Análisis de la covarianza					

FIG. 5.4 Clasificación de los diseños más importantes. (Tomado de M. Bartolomé, 1980)

DISEÑOS BIV ALENTES

Se utilizan cuando aplicamos dos niveles de la variable independiente y generalmente se componen de un grupo experimental y un grupo control, o bien de dos grupos experimentales (sin grupo control). Cuando se utiliza el grupo control es posible equilibrar la influencia de las variables extrañas, pues tenderán a influir con la misma intensidad en ambos grupos. De esta manera muchas variables extrañas pueden quedar neutralizadas.

Supongamos que pretendemos detectar el efecto relativo del tiempo de estudio (A) sobre la asimilación de un contenido (VD). Si queremos analizar el efecto de 2 horas de estudio (a), y pasamos una prueba previa para comprobar la homogeneidad o equivalencia inicial y conocer el nivel de la variable dependiente antes de aplicar la independiente, utilizaremos un diseño pretest-postest (tabla 5.16) con grupo control. A este grupo le asignaremos el nivel mínimo de la variable independiente.

El grupo control es un grupo de referencia, pues, al igual que en el experimental, las mismas variables extrañas contextuales influirán con la misma intensidad en ambos grupos. En consecuencia, las diferencias observadas al medir la variable dependiente en cada grupo podrán atribuirse al tratamiento aplicado, siempre que los grupos sean inicialmente equivalentes. El grupo control permite estimar los cambios debidos a la maduración $\bar{X}_1-\bar{X}_2$ en aquellos experimentos suficientemente prolongados.

TABLA 5.16 Diseño bivalente pretest - postest

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	$\bar{X}_1$	a_0	$\bar{X}_2$
2 Experimental	25	Azar	$\bar{X}_3$	a_1	$\bar{X}_4$

El pretest puede aplicarse a la variable dependiente o a una variable extraña con la que puede estar relacionada. Si omitimos la primera prueba y asumimos la equivalencia inicial de los grupos, el diseño se denomina *sólo postest* (tabla 5.17). Con esta modalidad podemos comparar los niveles alcanzados por la variable dependiente en el grupo control y en el experimental, pero no podemos estimar los cambios operados en el grupo control y en el experimental pomo poder comparar los niveles antes y después del tratamiento.

TABLA 5.17 Diseño bivalente sólo postest

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	-	a_0	$\bar{X}_1$
2 Experimental	25	Azar	-	a_1	$\bar{X}_2$

DISEÑOS MULTIV ALENTES

Se utiliza cuando se aplican tres o más niveles de la variable independiente (tabla 5.18).

Grupos	*n*	*Asignación*	*Pretest (VD)*	*Tratamiento (VI)*	*Postest (VD)*
1 Control	15	Azar	$\bar{X}_1$	a_0 (oh)	$\bar{X}_2$
2 Experimental	25	Azar	$\bar{X}_3$	a_1 (2h)	$\bar{X}_4$
3 Experimental	20	Azar	$\bar{X}_5$	a_2 (3h)	$\bar{X}_2$

Tanto en diseños bivalentes como multivalentes (pretest-postest), si se demuestra la equivalencia inicial (homogeneidad) de los grupos, bastará comparar las medias del

postest para comprobar el efecto relativo de las horas de estudio en el rendimiento (tabla 5.19).

TABLA 5.19 Posibles comparaciones que pueden efectuarse en diseños bivalentes y multivalentes

Objetivo	Comparación	Prueba
Homogeneidad o equivalencia inicial	$\bar{X}_1 - X_3 - X_5$	ANOVA
Homogeneidad final (prueba hipótesis, efecto VI)	$\bar{X}_2 \quad \bar{X}_4 \cdot \bar{X}_6$	ANOVA
Cambio en grupos experimentales	$\bar{X}_4 - \bar{X}_3$	t apareados
	$X_6 \cdot \bar{X}_5$	t apareados
Cambio en grupo de control (Influencia externa, madurez)	$\bar{X}_2 - \bar{X}_1$	t apareados

Si los grupos experimentales muestran un mayor cambio que el control puede inferirse que el tratamiento origina la diferencia.

Siempre que sea posible, es más adecuado utilizar diseños multivalentes, ya que los diseños bivalentes ofrecen una generalización muy restringida de las conclusiones que no serán extensibles a otros niveles de VI.

DISEÑOS FACTORIALES COMPLETAMENTE AL AZAR

Los diseños bivalentes y multivalentes constan sólo de una variable independiente con diferentes niveles, cada uno de los cuales se aplica a distinto grupo de sujetos. Dado que el investigador sólo está interesado en la acción de una variable independiente, se denominan *diseños simples* o *diseños unifactoriales.* En estos diseños se intenta mantener todo constante excepto la variable independiente, con el fin de poder atribuir a dicha variable los cambios observados en la variable dependiente. Sin embargo, se plantean dos dificultades:

1) Es casi imposible mantener constantes todas las variables excepto una. El control completo es muy difícil de conseguir.

2) En la realidad educativa es apropiado hablar de causación múltiple, pues en cada experimento una gran multiplicidad de variables afectan conjuntamente a la variable dependiente.

En consecuencia, más que tratar de hacer lo imposible para mantener constantes todas las variables extrañas (situación artificial y poco generalizable), podemos considerar la acción conjunta de dos o más variables independientes.

En un diseño multifactorial o «factorial», el investigador manipula dos o más variables independientes (cada una de ellas con dos o más niveles de tratamiento) con el fin de estudiar los cambios operados en la variable dependiente como consecuencia de dicha manipulación.

En los diseños factoriales se establecen todas las posibles combinaciones de los niveles de una variable independiente con los niveles de otra o más variables independientes. A cada combinación o casilla se asigna, al azar, un grupo de sujetos. Al diseño factorial de dos variables o factores se le denomina *bifactorial,* y si cada uno de ellos tiene dos niveles, se expresa mediante la notación 2 x 2, y exige formar 2 x 2 = 4 grupos experimentales o casillas. Un diseño bifactorial con tres niveles en cada factor se denomina 3 x 3. Así, un diseño 2 x 3 indica un experimento con dos factores, el primero con dos niveles y el segundo con tres. De la misma manera, un diseño 3 x 2 x 4 se refiere a la acción conjunta de tres factores o variables, la primera de las cuales actúa con tres niveles de tratamiento, la segunda con dos y la tercera con cuatro. Para aplicar este último diseño harían falta 3 x 2 x 4 = 24 grupos o casillas.

Aunque es casi ilimitado el número de variables que se pueden manipular, nos centraremos en diseños bifactoriales a dos niveles de acción (2 x 2), es decir, dos factores o variables independientes con dos niveles cada una. A partir de estos diseños se puede generalizar a otro tipo de diseños más complejos.

Veamos el siguiente ejemplo. Con 16 sujetos elegidos al azar se pretende comprobar cómo influyen en el rendimiento el método didáctico utilizado (pasivo o activo según el grado de participación del alumno) y la gratificación otorgada (alta o baja según la mayor o menor frecuencia de recompensas: bolígrafos, libros, pinturas, cromos, etc.) cada vez que el alumno asimila correctamente un contenido temático determinado. Podría establecerse una hipótesis como la siguiente: el tipo de método tiene un efecto diferente según se apliquen gratificaciones bajas o altas. Con el fin de facilitar la recogida de datos puede esquematizarse la matriz factorial (2 x 2) donde cada letra mayúscula representa una variable o factor, y las respectivas minúsculas, los niveles correspondientes. Los subíndices indican la magnitud de cada nivel del factor implicado, es decir, al < a2, bl < b2, 10 que da lugar a los siguientes símbolos:

		A	
		a,	a_2
B	b,	a, b,	a_2 b,
	b_2	a, b_2	a_2 b_2

Para llevar a cabo el experimento habría que formar cuatro grupos de sujetos al azar y asignarles una combinación de tratamientos también al azar. La representación simbólica del diseño 2 x 2 planteado también puede hacerse del siguiente modo:

Grupos	*n*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	4	Azar	-	a_1b_1	$\bar{X}_1$
2	4	Azar	-	a_1b_2	$\bar{X}_2$
3	4	Azar	-	a_2b_1	$\bar{X}_3$
4	4	Azar	-	a_2b_2	$\bar{X}_4$

Las respectivas medias de cada grupo se representarían del siguiente modo:

A
Método

B
Gratificación

	Pasivo	Activo
Baja	$\bar{X}_1$	$\bar{X}_3$
Alta	$\bar{X}_2$	$\bar{X}_4$

Las combinaciones de tratamientos, tal como han sido simbolizadas más arriba, son las siguientes:

A
Método

B
Gratificación

	Pasivo	Activo
Baja	Grupo 1	Grupo 3
Alta	Grupo 2	Grupo 4

Grupo 1: Las dos variables actúan con los niveles más bajos, es decir, método pasivo o basado en una participación mínima del alumno y en la poca adjudicación de materiales escolares como recompensa (gratificación baja).
Grupo 2: Los sujetos se someten al nivel más bajo de participación (método pasivo) mientras que la gratificación es máxima.
Grupo 3: Se concede a los alumnos el mayor grado de iniciativa y participación en las clases (método activo), pero el nivel de gratificación es el más bajo.
Grupo 4: Se combinan los niveles más altos de cada factor o variable: método activo y gratificación alta.

EFECTOS FACTORIALES

Los diseños factoriales aportan una información que no podría obtenerse en un diseño unifactorial de una sola variable independiente aunque ésta tenga más de dos niveles de tratamiento. Con un diseño factorial pueden analizarse los siguientes efectos:

a) *Efectos simples.* Expresan los efectos de los niveles de una variable independiente o factor dentro de cada nivel del otro factor. Así, en el ejemplo propuesto, el factor A (método) ejerce dos efectos simples. El primer efecto simple de A se obtiene al comparar las medias de A en el primer nivel del factor B (gratificación baja): X_3-X_1, El segundo efecto simple de A se obtiene al comparar las medias de A en el segundo nivel del factor B (gratificación alta): X_4-X_2. De la misma manera se pueden obtener los efectos simples del factor B (gratificación).

b) *Efectos principales o globales.* Se refieren a la influencia global de cada variable independiente por separado, es decir, sin tener en cuenta los valores que toma la otra variable independiente. El efecto principal equivale a la influencia ejercida cuando en el experimento interviene una sola variable independiente. En consecuencia, los efectos principales se pueden estimar calculando las diferencias entre las medias de las filas y entre las medias de las columnas. Ahora bien, dado que los efectos simples son estimaciones del efecto principal, para calcular el efecto principal también pueden promediarse los efectos simples.

c) *Efectos secundarios o de interacción.* Existe interacción entre dos variables independientes cuando el valor que toma la variable dependiente, como resultado de las modificaciones introducidas en aquéllas, depende de los valores que toma la otra variable independiente. La interacción, por tanto, supone que un tratamiento tiene diferentes efectos al combinado con diferentes niveles de otro tratamiento. En definitiva, la influencia simultánea de varios factores hace que ciertas combinaciones entre ellos sean más efectivas que otras.

En el ejemplo propuesto, gracias al diseño factorial podemos comprobar si el efecto del método activo (o del método pasivo) varía a medida que se aplica combinado con los respectivos niveles de la variable gratificación (alta y baja). Si la variación fuera efectiva (significativa) es que existe interacción entre las dos variables o factores (método y gratificación).

La interacción es significativa cuando la diferencia entre los efectos simples de un factor es demasiado grande para ser atribuible al azar. Así, la comparación de los efectos simples para un determinado factor es una estimación de la posible existencia de interacción. En un diseño factorial 2 x 2 la interacción también puede estimarse calculando la diferencia entre la suma de los valores de las diagonales en la tabla de datos promedio.

Para determinar si existe interacción entre dos o más variables puede procederse de dos maneras:

1) *Análisis de la varianza.* Proporciona información sobre la significación estadística de la interacción. Su estudio pormenorizado se expondrá en capítulos posteriores al tratar el análisis de datos.

2) Representación gráfica. La inspección visual de la gráfica permite apreciar de forma intuitiva si existe o no interacción. En el eje de abscisas se representan los valores de uno de los factores y en el de ordenadas se representan los valores de la variable dependiente. Al unir los puntos correspondientes a los niveles del segundo factor, las rectas resultantes pueden tomar alguna de las siguientes formas:

- rectas paralelas;
- rectas que no son paralelas, aunque no llegan a cortarse en la gráfica;
- rectas que se cortan en la gráfica.

Ilustraremos la interpretación gráfica de la interacción a partir del ejemplo ya planteado para comprobar cómo influyen en el rendimiento el método didáctico utilizado (pasivo-activo) y la gratificación (alta-baja). También se describirán las tres situaciones contemplando distintos tipos de efectos factoriales como son: los efectos principales o globales, secundarios o de interacción, y simples.

PRIMERA SITUACIÓN: RECTAS PARALELAS

Supongamos que en el ejemplo planteado las puntuaciones de rendimiento obtenidas por los 16 sujetos en el postest fueran las siguientes:

Sujeto	*Tratamiento*	*Rendimiento*
1	a_1b_1	0
2	a_1b_1	1
3	a_1b_1	2
4	a_1b_1	1
5	a_1b_2	3
6	a_1b_2	4
7	a_1b_2	3
8	a_1b_2	2

Sujeto	*Tratamiento*	*Rendimiento*
9	a_2b_1	4
10	a_2b_1	5
11	a_2b_1	6
12	a_2b_1	5
13	a_2b_2	8
14	a_2b_2	6
15	a_2b_2	8
16	a_2b_2	6

Estos datos pueden tabularse de distintas maneras. Cuando queremos analizar los efectos globales de cada variable puede ser útil una tabla como la siguiente:

		A: a_1	A: a_2	$\bar{X}$
B	b_1	0 1 2 1	4 5 6 5	3
B	b_2	3 4 3 2	8 6 8 6	5
	$\bar{X}$	2	6	4

Como ejemplo ilustrativo se han registrado todas las puntuaciones directas de rendimiento; sin embargo, presenta mayor utilidad incluir las medias obtenidas por cada grupo en una tabla de datos promedio como se verá más adelante.

Para apreciar con más claridad el efecto del método en cada nivel de gratificación también puede elaborarse una tabla como la siguiente:

Método	*Pasivo*		*Activo*	
Gratificación	Baja	Alta	Baja	Alta
Puntuaciones de rendimiento	1 2 1	4 3 2	5 6 5	6 8 6
Rendimiento medio	1	3	5	7

A partir de las tablas anteriores pueden estimarse los efectos factoriales principales o globales, secundarios o interacciones, y simples o diferenciadores.

a) *Efectos factoriales principales o globales*

Se refieren a la influencia global de las variables método y gratificación consideradas por separado. En definitiva, es como si hubiéramos realizado dos experimentos bivalentes, uno para ver qué método es más efectivo (el activo o el pasivo) y otro para ver qué gratificación es más efectiva (la alta o la baja). La comparación de los resultados finales de cada columna nos permite comprobar si los sujetos rinden más con un método pasivo que con un método activo. Al comparar los resultados de las filas podemos estimar la eficacia relativa de cada tipo de gratificación (baja o alta):

1) Efecto global de la gratificación. El efecto principal de la gratificación vale 5 - 3 = 2. Es decir, al pasar de un nivel de gratificación bajo a un nivel alto el rendimiento aumenta en 2 unidades:

Gratificación	*Baja*	*Alta*
	0	3
	1	4
	2	3
Puntuaciones	1	2
de rendimiento	4	8
	5	6
	6	8
	5	6
Rendimiento medio	3	5

2) Efecto global del método. El efecto principal del método vale 6- 2 = 4. Es decir, al pasar de un método pasivo a un método activo el rendimiento aumenta en cuatro unidades:

Método	*Pasivo*	*Activo*
	0	4
	1	5
	2	6
Puntuaciones	1	5
de rendimiento	3	8
	4	6
	3	8
	2	6
Rendimiento medio	2	6

b) *Efecto factorial secundario o interacción*

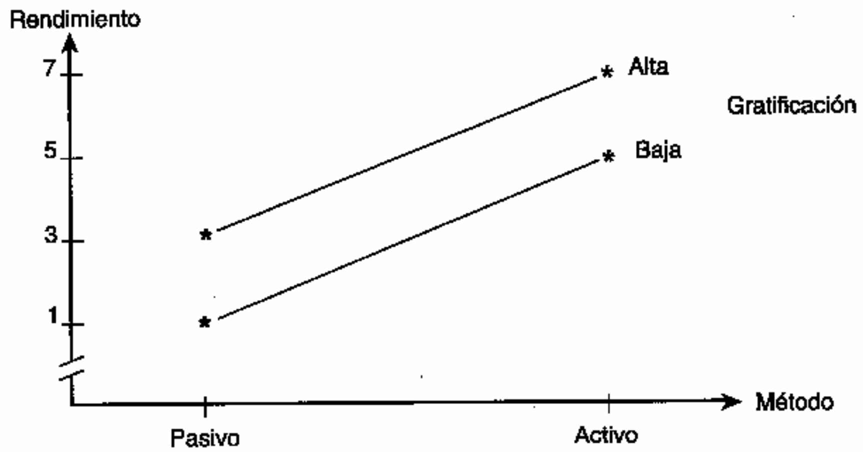
Las conclusiones anteriores de que la gratificación alta es más efectiva que la baja porque aumenta el rendimiento en dos unidades (suponiendo que hubiera significación estadística) y de que el método activo es más efectivo que el pasivo tienen sentido en el caso de que no existan efectos factoriales secundarios o interacciones.

No existe interacción entre la gratificación y el método cuando el aumento del rendimiento provocado por la gratificación alta (dos unidades) es el mismo al combinarse con el método pasivo y con el activo. Es decir, al pasar de una gratificación baja a otra alta el rendimiento aumenta una cantidad constante que siempre vale dos unidades, tanto si el método es pasivo como si es activo. Esto significa que los efectos simples correspondientes a los niveles de una misma variable o factor son estadísticamente iguales. La tabla de datos promedio es la siguiente:

Método

Gratificación	Pasivo	Activo	
Baja	1	5	3
Alta	3	7	5
	2	6	4

Puede comprobarse que al pasar de un método pasivo a un método activo el rendimiento también aumenta una cantidad constante de cuatro unidades, tanto para la gratificación baja como para la gratificación alta. Es decir, los efectos simples vuelven a ser iguales. La ausencia de interacción o igualdad de efectos simples queda reflejada por la presencia de dos rectas paralelas:



La interacción puede estimarse a partir de la diferencia entre la suma de los valores de las diagonales:

$$[(1 + 7) - (3 + 5)] /2 = O.$$

Cuando la interacción no es significativa se puede concluir que los efectos factoriales son «aditivos» o que existe aditividad, y por tanto, que existe independencia entre los factores o variables. Aunque la interacción no sea significativa, en algunos casos el análisis de los efectos simples puede ser muy revelador.

c) *Efectos factoriales simples*

Si nos fijamos en la variable gratificación, podemos determinar dos efectos de la misma:

1) Efecto simple de la gratificación cuando el método es pasivo: 3- 1 = 2.
2) Efecto simple de la gratificación cuando el método es activo: 7- 5 = 2.

Esta información indica que el efecto de la gratificación tiene la misma intensidad al combinarse con ambos métodos. Con respecto a la variable método puede procederse de la misma manera:

3) Efecto simple del método cuando la gratificación es alta: 7- 3 = 4.
4) Efecto simple del método cuando la gratificación es baja: 5- 1 = 4.

Estos resultados indican que el método tiene una influencia idéntica cuando la gratificación es baja y cuando es alta. Cuando las diferencias entre los efectos factoriales simples de una variable pueden atribuirse estadísticamente al azar no existen efectos factoriales secundarios o interacción. Así, para la variable gratificación, sus efectos simples (2 y 2) son dos estimaciones de la influencia de dicha variable. Si son estadísticamente iguales, hay que concluir que la variable método no interviene en ellos y que no existe interacción. Es decir, la variable gratificación, al combinarse con los dos tipos de método, influye por igual en el rendimiento.

Puede comprobarse que el efecto principal o influencia global de cada variable puede estimarse también a partir de los efectos simples:

- Influencia global de la gratificación: promedio de los efectos simples (2 + 2) / 2 = 2. Como ya hemos visto, dicho efecto se puede calcular recurriendo a la diferencia entre las medias de las filas: 5- 3 = 2.
- Influencia global del método: promedio de los efectos simples (4 + 4) / 2 = 4.

El mismo efecto se puede calcular con la diferencia entre las medias de las columnas: 6-2 = 4.

También puede estimarse la interacción a partir de la diferencia entre los efectos simples, ya sea 2- 2 = 0, o bien, 4- 4 = 0.

SEGUNDA SITUACIÓN: LAS RECTAS NO SON PARALELAS PERO NO SE CORTAN EN LA GRÁFICA

También podríamos haber obtenido datos ficticios como los siguientes:

		Método		
		Pasivo	Activo	
Gratificación	Baja	3	1	2
	Alta	5	7	6
		4	4	4

Según la tabla, los tipos de efectos son:

a) *Efectos factoriales principales* o *globales*

1) Influencia global de la gratificación: 6 - 2 = 4. Estadísticamente podría existir diferencia significativa y ser más efectivo el nivel alto de gratificación.
2) Influencia global del método: 4- 4 = O. El método activo tiene la misma efectividad que el método pasivo, ambos tienen un efecto idéntico.

b) *Efecto factorial secundario* o *interacción*

Las conclusiones anteriores son globales y pueden inducir a error. No son válidas por igual cuando cada variable se combina con los niveles de la otra. Según los efectos globales de la gratificación, al pasar de un nivel bajo de gratificación a un nivel alto aparece un aumento en rendimiento de cuatro unidades. Si este aumento fuera estadísticamente igual (cuatro unidades) bajo los dos métodos (activo y pasivo), no existiría interacción y el efecto global sería válido. Sin embargo, bajo el método pasivo el aumento es de sólo dos unidades, pero bajo el método activo el aumento es de seis unidades, lo que origina una aumento global medio de (2 + 6) / 2 = 4 unidades, como ya se había calculado. Este aumento es debido al cambio introducido en los niveles de gratificación.

Si sólo nos basamos en los efectos globales del método, el error puede ser aún más grave. Según los efectos principales o globales, tanto el método pasivo como el activo tienen la misma efectividad (4 - 4 = O). Ahora bien, cuando la gratificación es baja, al pasar de un método pasivo a un método activo el rendimiento sufre un descenso de dos unidades (1- 3 = -2). En cambio, cuando la gratificación es alta, aparece un aumento de 7 - 5 = 2 unidades; de ahí que se detecte un aumento global medio que vale

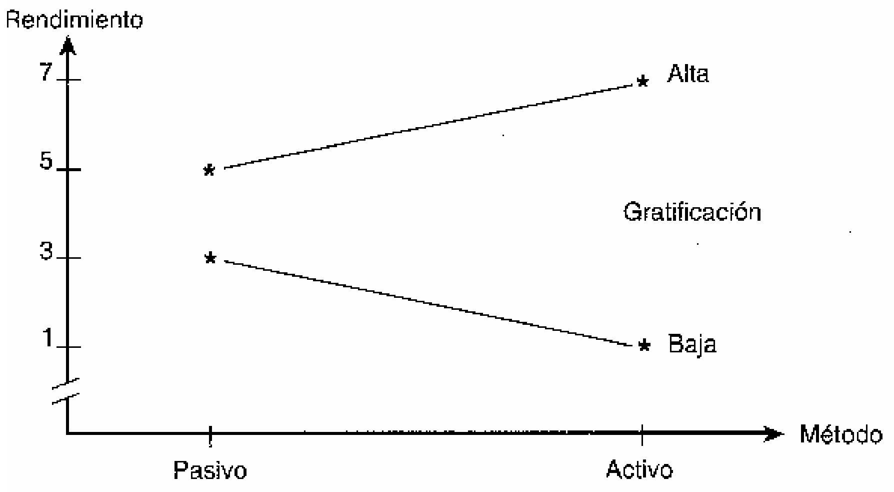
[2 + (- 2)] /2 = O.

Es decir, a pesar de que ambos métodos tienen el mismo efecto global, sus efectos factoriales simples son diferentes para cada nivel de gratificación.

Se produce una interacción entre las variables método y gratificación cuando los cambios observados en el rendimiento como consecuencia del método son significativamente diferentes para cada tipo de gratificación; es decir, la efectividad del método es distinta según el nivel de gratificación, y viceversa. El cálculo numérico de la interacción sería

$$[(3 + 7) - (5 + 1)] /2= 2$$

Al existir interacción no es suficiente analizar los efectos globales o principales de cada variable; es imprescindible analizar los efectos simples. Cuando existe interacción las rectas no tienden a ser paralelas:



c) *Efectos factoriales simples*

Con respecto a la variable gratificación hemos estimado dos efectos de la misma:

1) Efecto factorial simple de la gratificación cuando el método es pasivo: 5-3 = 2.
2) Efecto factorial simple de la gratificación cuando el método es activo: 7-1 = 6.

En consecuencia, el efecto de la gratificación tiene diferente intensidad (debería comprobarse si la diferencia es excesiva para ser atribuida sólo al azar) cuando el método es pasivo y cuando es activo.

Con respecto a la variable método puede procederse de la misma manera:

3) Efecto factorial simple del método cuando la gratificación es alta: 7-5 = 2.
4) Efecto factorial simple del método cuando la gratificación es baja: 1-3 = -2.

Como puede apreciarse, el efecto es negativo porque el nivel más participativo del método (activo) es menos efectivo que el nivel más bajo de participación (pasivo).

Estos resultados indican que el método tiene una influencia distinta según sea el nivel de gratificación (alta o baja). Cuando la gratificación es alta, al pasar de un método pasivo a un método activo se aprecia que el rendimiento aumenta en dos unidades. En cambio, cuando la gratificación es baja, ocurre al contrario: al pasar de un método pasivo a un método activo el rendimiento disminuye en dos unidades. Como es evidente, los efectos globales pueden calcularse recurriendo a los efectos globales:

- Influencia global de la gratificación: promedio de los efectos simples (2 + 6) *12* = 4. El mismo efecto se puede calcular a partir de la diferencia entre las medias de las filas: 6 - 2 = 4. - Influencia global del método: promedio de los efectos simples [2 + (- 2)] *12* = 0.

De la misma manera, a partir de la diferencia entre las medias de las columnas: 4 - 4=0.

TERCERA SITUACIÓN: LAS RECTAS SE CORTAN EN LA GRÁFICA

Otra posibilidad es que los datos resultantes sean como los de la tabla que se adjunta:

a) *Efectos factoriales principales o globales*

1) Influencia global de la gratificación: 5 - 5 = 0. Los dos niveles de gratificación tienen la misma efectividad.

2) Influencia global del método: 6- 4 = 2. El método activo aumenta dos unidades el rendimiento con respecto al método pasivo.

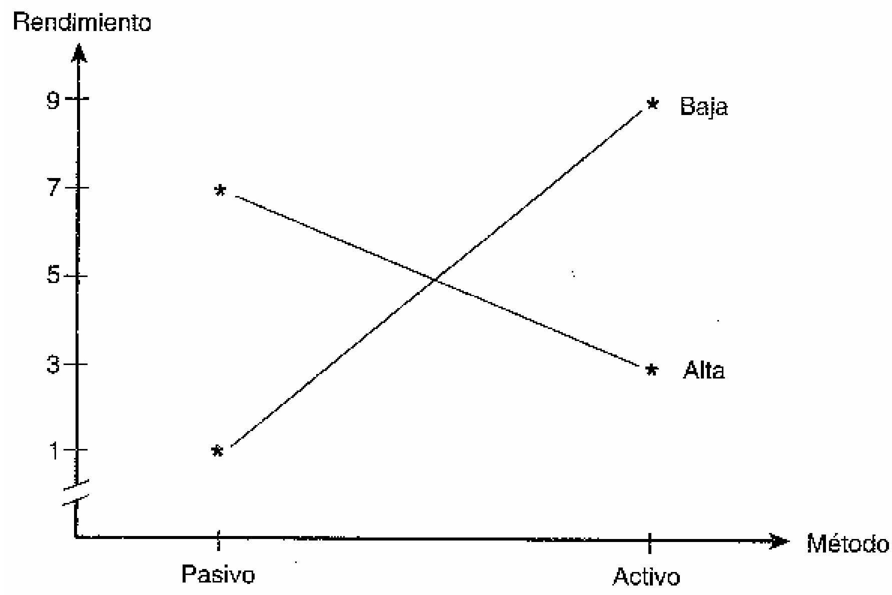
Método

		Pasivo	Activo	
Gratificación	Baja	1	9	5
	Alta	7	3	5
		4	6	5

b) *Efecto factorial secundario a interacción*

Cuando en la gráfica aparecen rectas que se cortan es muy posible que exista interacción entre los dos factores (gratificación y método), cuyo valor sería:

[(1 + 3) - (7 + 9)] *12* = - 6



Cuando exista interacción es necesario analizar los efectos simples.

c) *Efectos factoriales simples*

1) Efecto factorial simple de la gratificación cuando el método es pasivo: 7 - 1 = 6.
2) Efecto factorial simple de la gratificación cuando el método es activo: 3 - 9 = - 6.

Con respecto a la variable método puede procederse de la misma manera:

3) Efecto factorial simple del método cuando la gratificación es alta: 3 - 7 = - 4.
4) Efecto factorial simple del método cuando la gratificación es baja: 9 - 1 = 8.

Para comprobar los efectos globales:

- Influencia global de la gratificación: promedio de los efectos simples:

$$[6 + (- 6)] / 2 = 0, \text{ o bien } 5 - 5 = 0.$$

- Efecto global del método: [8 + (- 4)] /2 = 2, o bien, 6 - 4 = 2.

5.7 DISEÑOS DE BLOQUES HOMOGÉNEOS AL AZAR

En los diseños completamente al azar la heterogeneidad entre los sujetos puede originar una varianza intragrupo (o del error) muy elevada. Con el fin de reducir la variabilidad de los sujetos y hacer más sensible el experimento se puede introducir la

técnica del bloqueo. De esta manera se pueden apreciar mejor los efectos de VI. Supongamos que estamos interesados en ver el efecto de tres niveles de gratificación (A = alto, M = medio y B = bajo) en el rendimiento de 12 niños de primaria. Para realizar un experimento de bloques hay que tener en cuenta las siguientes fases:

1) Mídase alguna variable extraña que sospechemos que pueda estar estrechamente relacionada con la variable dependiente. En consecuencia, podría contaminar el experimento. Dicha variable se denomina *variable de bloqueo* o, en su caso, *de apareo* (cuando cada bloque está formado por dos sujetos). Generalmente, dicha variable no coincide con la variable dependiente. En el ejemplo planteado consideramos que la variable inteligencia sería una variable de bloqueo adecuada. Después de medir la inteligencia obtenemos las siguientes puntuaciones:

Sujetos	*Inteligencia*
Antonio	98
Rosa	100
Carlos	94
Ignacio	94
Gema	96
Maribel	98

Sujetos	*Inteligencia*
Mónica	96
José	100
Laura	94
Ana	98
Nuria	100
Raquel	96

2) En función de las puntuaciones obtenidas en la variable de bloqueo, fórmense grupos de sujetos (bloques) con puntuaciones iguales o lo más similares posible. Dentro de cada bloque se asignan al azar, como mínimo, tantos sujetos o unidades experimentales como niveles tenga la variable independiente:

Bloques	*CI*		*Sujetos*	
Bloque 1	94	Ignacio	Carlos	Laura
Bloque 2	96	Gema	Mónica	Raquel
Bloque 3	98	Maribel	Antonio	Ana
Bloque 4	100	Nuria	José	Rosa

3) Asígnese cada nivel de la variable independiente al azar:

			Gratificación	
Bloques	*CI*	*Baja*	*Media*	*Alta*
Bloque 1	94	Ignacio	Carlos	Laura
Bloque 2	96	Gema	Mónica	Raquel
Bloque 3	98	Maribel	Antonio	Ana
Bloque 4	100	Nuria	José	Rosa

4) Aplíquense los niveles de la variable independiente.
5) Mídase la variable dependiente:

		Gratificación		
Bloques	el	*Baja*	*Media*	*Alta*
Bloque 1	94	9	9	9
Bloque 2	96	8	7	6
Bloque 3	98	6	5	4
Bloque 4	100	4	4	2

Si cada bloque está formado por dos sujetos, puede tratarse de un diseño de grupos apareados. Cuando cada unidad experimental está formada por grupos de sujetos, la variable de bloqueo puede considerarse como otra variable independiente o factor, y el experimento se convierte en factorial.

El diseño planteado se denomina *de bloques al azar* 3 4 porque el factor manipulado adopta tres niveles y se han formado cuatro bloques de sujetos con la variable de bloqueo. La representación simbólica del diseño sería la siguiente:

Grupos	*n*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	3	BA_4	-	a_1	$\bar{Y}_1$
2	3	BA_4	-	a_2	$\bar{Y}_2$
3	3	BA_4	-	a_3	$\bar{Y}_3$
4	3	BA_4	-	a_4	$\bar{Y}_4$

La técnica del bloqueo adolece de tres limitaciones importantes. En primer lugar, cuando se ha elegido un sujeto o grupo de sujetos para formar un bloque, el investigador debe seleccionar más sujetos semejantes entre los restantes para completar el bloque iniciado. Con frecuencia es difícil encontrar sujetos suficientes. Esta circunstancia origina la segunda limitación: los sujetos que no son similares a los incluidos en los bloques deben excluirse, con lo que disminuye la representatividad de la muestra con respecto a la población. Por último, la homogeneidad entre los sujetos de un mismo bloque se establece con respecto a la variable de bloqueo, pero los sujetos incluidos en un mismo bloque pueden diferir en alguna otra variable.

5.8 DISEÑOS INTRAGRUPO O DE MEDIDAS REPETIDAS

La eficacia y sensibilidad de un experimento aumenta en proporción a la homogeneidad de los sujetos, ya que, con grupos experimentales más homogéneos, se tiene mayor probabilidad de comprobar, en el caso de que exista, la posible influencia

de la variable independiente sobre la variable dependiente. La varianza del error disminuye progresivamente al pasar de los diseños completamente al azar a los diseños de bloques homogéneos. Con los diseños intragrupo la técnica del bloqueo se lleva hasta sus últimas consecuencias. Si a cada sujeto se le aplican todos los niveles de la variable independiente y se comparan entre sí los efectos producidos, hemos llevado hasta el extremo las condiciones de homogeneidad con respecto a las diferencias individuales. De hecho, cada sujeto puede considerarse como un solo bloque.

En los diseños intragrupo es importante el orden de aplicación. Los resultados obtenidos en la variable dependiente pueden depender de la secuencia seguida en la aplicación de los tratamientos. Para evitarlo es viable aleatorizar el orden de aplicación o puede procederse a un contrabalanceo de los tratamientos.

La representación simbólica de un diseño intragrupo puede realizarse del siguiente modo:

Grupos	*número de sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n	A	-	$a_1; a_2; a_3$	$\bar{X}_1; \bar{X}_2; \bar{X}_3$

Veremos dos aplicaciones de los diseños intragrupo: el diseño tratamiento por sujetos y el análisis de tendencias.

DISEÑO DE TRATAMIENTOS POR SUJETOS

Cuando estamos interesados en comparar la efectividad de una serie de tratamientos o niveles de la variable independiente se utiliza el denominado *diseño de tratamientos por sujetos.*

La representación simbólica anterior puede hacerse extensible del siguiente modo:

Grupos	*Pretest (VI)*	*Tratamiento (VD)*	*Postest*
1	-	$a_1 \ a_2 \ a_3 \ldots a_k$	$X_1 \ X_2 \ X_3 \ldots X_k$
1	-	$a_1 \ a_2 \ a_3 \ldots a_k$	$X_1 \ X_2 \ X_3 \ldots X_k$
1	-	$a_1 \ a_2 \ a_3 \ldots a_k$	$X_1 \ X_2 \ X_3 \ldots X_k$
•			
•			
•			
n	-	$a_1 \ a_2 \ a_3 \ldots a_k$	$X_1 \ X_2 \ X_3 \ldots X_k$
			$\bar{X}_1 \ \bar{X}_2 \ \bar{X}_3 \ldots \bar{X}_k$

En la medida de lo posible, el orden de aplicación de los niveles de la variable independiente será totalmente aleatorio. La estructura lógica de este diseño es similar a la del diseño de bloques, ya que cada sujeto equivale a un bloque.

ANÁLISIS DE TENDENCIAS

Se utiliza cuando el énfasis se centra en la mejora o cambio que se produce como consecuencia de una práctica más o menos prolongada. Los niveles de la variable independiente difieren entre sí según una cantidad constante. Se trata de comprobar si las medias obtenidas en la variable dependiente bajo cada tratamiento tienden a variar en un sentido lineal, curvilíneo, etc., y hasta qué punto es significativa esta tendencia. Es un diseño muy común en situaciones de aprendizaje donde interesan los efectos acumulados de los tratamientos (fig. 5.5).

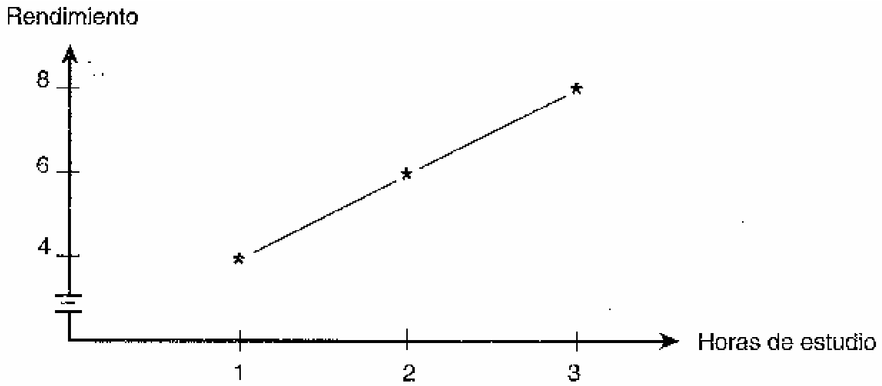


FIG. 5.5 Función lineal de las medias de rendimiento según las horas de estudio en un experimento intragrupo

5.9 TIPOS DE EXPERIMENTOS

Para conseguir el máximo grado de control posible en los experimentos se tiende a un aislamiento del contexto natural donde se produce el fenómeno. La finalidad de este aislamiento radica en el objetivo de eliminar las múltiples influencias extrañas capaces de actuar juntamente con la variable independiente, pues la acción de dichas variables la confundiríamos con la influencia de la variable independiente sobre la variable dependiente.

En la medida que controlamos las variables extrañas (restricciones) garantizamos la validez interna, pero las condiciones en que se desarrolla el experimento son más artificiales y, por tanto, menos representativas, pues en un contexto real las variables no se hallan tan controladas. En consecuencia, la generalización a situaciones reales queda reducida. De ahí la conveniencia de diferenciar entre experimentos de laboratorio y de campo, teniendo en cuenta que entre ambas modalidades sólo existe una diferencia de grado, los primeros con un control máximo, y los segundos, con un control menor.

Experimento de laboratorio. Es un estudio de investigación donde se consigue un aislamiento del contexto natural en que se produce el fenómeno con el fin de eliminar las múltiples influencias extrañas capaces de afectar a la variable dependiente.

Experimento de campo. Es un estudio de investigación en una situación real donde una o más variables independientes son manipuladas por el experimentador en condiciones controladas con el máximo rigor que permita la situación.

5.10 POSIBILIDADES Y LÍMITES

El método experimental se aplica a través de diseños específicos que suponen adoptar variaciones respecto a la categoría global, y que posibilitan responder de forma más acorde a determinados problemas y situaciones de investigación. En cuanto a sus aplicaciones a cuestiones o problemas educativos, los diseños más rigurosos se han utilizado principalmente en el área de tecnología educativa y didáctica, configurando la inmensa mayoría de trabajos que se han publicado.

a) *Posibilidades.* Según algunos autores (Mayntz, 1983), es el más prominente de todos los métodos de investigación en cuanto que es el único que permite constatar estrictamente relaciones causales. Al manipular o provocar el fenómeno, el investigador está en condiciones idóneas para inferir proposiciones causales (Cook y Campbell, 1979; Arnau, 1978; Haimson y Elfebein, 1985). Otras funciones básicas del método experimental son descubrir relaciones en condiciones puras y no contaminadas, probar predicciones derivadas de la teoría o de otra investigación y desarrollar teorías e hipótesis para elaborar sistemas teóricos.

Estas funciones son viables gracias al elevado grado de control que se ejerce sobre:

- Las variables independientes, mediante su modificación en cantidad, grado o cualidad.
- Las variables extrañas, cuyo dominio o control facilitará la maximización de la varianza primaria.
- La medición, que deberá ser fiable, precisa y válida.

El control relativamente completo permite una gran especificidad en las definiciones operativas de las variables, con lo que se favorece la precisión y la posibilidad de llevar a cabo las réplicas necesarias. La investigación queda muy bien delimitada, es inequívoca y los errores de medida son mínimos, ya que las condiciones del experimento se especifican minuciosamente. Como consecuencia, la validez interna es alta, con lo que estamos ante un enfoque adecuado para probar diferentes aspectos de la teoría.

Otra característica vinculada a la necesidad de inferir efectos de los tratamientos es la utilización de la comparación. Existen muchos tipos de comparación en la experimentación que responden a diferentes propósitos y originan distintas modalidades de diseños de investigación. También cabe destacar su precisión y sensibilidad a la hora de maximizar la varianza experimental y minimizar la del error, como consecuencia del control de las situaciones en las que se produce el experimento (Kerlinger, 1982). En resumen, las ventajas y posibilidades de esta metodología se pueden resumir en precisión y economía (Mouly, 1978).

Los *experimentos de laboratorio* son especialmente idóneos para cubrir objetivos de investigación como los siguientes:

- Constatar relaciones en condiciones «puras» e incontaminadas.
- Someter a prueba las predicciones derivadas de la teoría (hipótesis).
- Generar nuevas investigaciones.
- Refinar las teorías y las hipótesis. Ayudan a construir sistemas teóricos.

El experimento de laboratorio presenta ventajas como las siguientes:

- Control relativamente completo.
- Permite la asignación aleatoria y la manipulación de una o más variables independientes.
- Alto grado de especificidad en las definiciones operativas de las variables.
- Precisión. Bien delimitado e inequívoco. La varianza del error es pequeña, ya que las condiciones del experimento se especifican minuciosamente disminuyendo los errores.
- Alta validez interna.
- Flexibilidad: hay muchas posibilidades experimentales y se pueden poner a prueba variados aspectos de la teoría.
- Replicabilidad: posibilidad de repetir el experimento introduciendo pequeñas variantes.

Las ventajas principales de los experimentos de campo pueden ser las siguientes:

- Muy utilizados en pedagogía, se llevan a cabo en aulas, colegios y otras instituciones y comunidades. Las variables actúan en su propio medio natural. - Al menos en teoría permiten manipular las VI y asignar sujetos y tratamientos al azar.
- Las variables suelen tener un efecto más potente que las de laboratorio. Cuanto más realista es la situación más potentes serán las variables.
- Adecuados para estudiar las influencias sociales complejas, los procesos y cambios sociales en situaciones de la vida real.
- Pueden utilizarse para someter a prueba las teorías y para solucionar problemas prácticos.
- Flexibilidad y aplicabilidad a una gran variedad de problemas.

b) *Límites.* Pero los experimentos no están exentos de algunos inconvenientes.

Como contrapartida existen limitaciones al aplicar esta metodología al campo pedagógico, como son la selección de las variables, la naturaleza re activa de los sujetos, posibles sesgo s introducidos por el experimentador, limitaciones procedentes del muestreo, problemas deontológicos (Pérez Juste, *1985a;* Mialaret, 1972), además de los de generalización como consecuencia de la artificialidad del control (Cook y Campbell, 1979).

Así, dada la artificialidad de la situación experimental, puede ocurrir que el influjo de la variable independiente sobre la dependiente sea débil, ya que la situación se ha provocado con propósitos especiales. Por ello, hay que tener en cuenta que los experimentos, especialmente los de laboratorio, presentan poca generalización. Esta circunstancia disminuye su validez externa, con el consiguiente riesgo de interpretar o extrapolar erróneamente los resultados.

Algunas de las limitaciones comentadas deben matizarse si se tiene en cuenta que los experimentos, más que generalizar, pretenden constatar relaciones de causalidad en condiciones puras y no contaminadas, sometiendo a prueba las predicciones derivadas de la teoría. De ahí que **un** cometido fundamental del enfoque experimental sea refinar las teorías a través del contraste de hipótesis, con lo que se favorece la construcción de sistemas teóricos.

Aunque parece ser el método idóneo para garantizar la validez interna, algunos autores (Harsé y Secord, 1976) la cuestionan a la hora de estudiar la conducta social, señalando limitaciones como las que impone la misma naturaleza de la situación, la operacionalización de conceptos, el desconocimiento de parámetros relevantes y el tipo de interacción peculiar de los experimentos.

Algunas limitaciones de los experimentos de laboratorio son las siguientes:

- Debilidad de las variables independientes, ya que se han creado las situaciones con propósitos especiales.
- Artificialidad de la situación experimental si no se tienen en cuenta los objetivos de los experimentos de laboratorio (relaciones de causación entre variables). - Baja validez externa, poca generalización, pero su objetivo primordial no es generalizar.

En cuanto a los experimentos de campo, pueden destacarse las siguientes limitaciones:

- Carecen de **un** control riguroso.
- Objeciones ante la manipulación de las variables que puedan afectar a la educación o al rendimiento del alumno.
- Falta de precisión.
- Requieren en cierta medida **un** experto en trabajo social.

CAPÍTULO 6

Metodología cuasiexperimental

En las investigaciones cuasiexperimentales, el investigador varía deliberadamente los niveles de la variable independiente para poder ver los efectos que causa dicha variación en la variable dependiente, pero no ejerce el grado de control característico del método experimental. Muchas variables extrañas quedan sin controlar. En el experimento, la muestra se elige al azar de la población, los grupos se forman al azar y los distintos niveles de la variable independiente a los que se exponen los diferentes grupos se asignan a éstos al azar. En la investigación cuasiexperimental puede faltar alguno o varios de estos requisitos.

Generalmente, la metodología cuasiexperimental se lleva a cabo en una situación real o de campo, donde una o más variables independientes son manipuladas por el investigador en condiciones controladas sólo hasta donde permita la situación. Efectivamente, en muchas situaciones educativas el investigador encuentra obstáculos para ejercer el grado de control que requieren los experimentos estrictos, por lo que habrá de tener en cuenta que algunas variables han quedado sin controlar. Por tanto, existirá la posibilidad de que la variación observada en la variable dependiente se deba más a la acción de tales variables que a la del factor manipulado. En tales situaciones se considera que la investigación tiene un carácter cuasiexperimental y suele emplearse en contextos educativos donde no es viable alterar la estructura o configuración de grupos ya formados, con lo que es difícil poder aleatorizar los sujetos e incluso los tratamientos.

6.1 TIPOS DE DISEÑOS

En general, la metodología cuasiexperimental pretende explicar relaciones de causalidad comparando grupos de datos procedentes de situaciones provocadas por el investigador pero que carecen de un control completo.

Los diseños *cuasiexperimentales* más empleados pueden agruparse en tres categorías globales (fig. 6.1).

a) *Diseños de grupos no equivalentes.* Se utilizan con mucha frecuencia en el ámbito educativo. Hay que optar por esta vía metodológica cuando el investigador

pretende analizar relaciones de causalidad y puede manipular la variable independiente, pero, a diferencia de la metodología experimental, se ve obligado a partir de grupos ya formados de una manera natural, como pueden ser las clases de un colegio.

Los distintos diseños se originan al combinar el número de grupos que intervienen y la posibilidad de aplicar pretest-postest o un sólo postest. A modo de ejemplos ilustrativos nos centraremos en diseños de grupo único y de dos grupos, con la doble modalidad pretest-postest o sólo postest.

Grupos no equivalentes	*Series temporales interrumpidas*	*Sujeto único*
Grupo único sólo postest	Diseño simple	Diseño AS
Grupo único pretest-postest	Dos grupos no equivalentes	Diseño ASA
Dos grupos sólo postest	Retiradade tratamiento	Línea base múltiple
Dos grupos pretest-postest	Replicaciones múltiples	

FIG. 6. 1 Tipos de diseño cuasiexperimentales

b) *Diseños de series temporales interrumpidas.* El método basado en diseños de series temporales está más indicado cuando el problema que hay que investigar exige tomar una serie de medidas de la variable dependiente a lo largo de **un** determinado período, interrumpiendo la serie con la aplicación de algún tratamiento (Cook y Campbell, 1979). Con posterioridad a la introducción del cambio se sigue tomando otra serie de medidas de la variable dependiente con el fin de poder detectar si la secuencia sufre cambios de nivel y/o tendencia (Pereda, 1987). Estos diseños pueden plantear inconvenientes al ser aplicados en algunos sectores del ámbito educativo, ya que ciertos cambios o tratamientos deben introducirse de forma gradual y los efectos pueden ser muy retardados y difíciles de detectar. Por otro lado, es imprescindible recoger datos durante un período tanto más amplio cuanto más irregular es la secuencia.

Los diseños de series temporales pueden ser de distintos tipos. Cuando interviene un solo grupo de sujetos, el diseño se denomina *simple,* pero cuando puede recurrir a otro grupo, que suele ser un grupo control, el diseño se denomina *de dos grupos no equivalentes.* Cuando queremos tener más confianza en las conclusiones podemos emplear un diseño con retirada de tratamiento, lo que permite comparar tres períodos. Por último, cuándo la situación anterior se repite sucesivamente origina un diseño con *replicaciones múltiples.*

c) *Diseños de sujeto único.* En general, la investigación basada en la metodología empírico- analítica o cuantitativa enfatiza la generalización de los resultados, y por ello tiende a centrarse en investigaciones cuyos diseños implican grupos de sujetos. Los denominados diseños de N = 1 o de sujeto único o intrasujeto constituyen una alternativa a la estrategia basada en grupos (diseños intergrupos, intragrupos y mixtos). Las aplicaciones de los diseños implicados son muy variadas y actualmente

se utilizan frecuentemente en el ámbito del condicionamiento operante (modificación de conducta), conducta disocia1, trastornos clínicos, aplicaciones terapéuticas, intervenciones educativas, retraso mental y fracaso escolar. No hay ninguna razón que justifique el uso restringido que han tenido los diseños de sujeto único, ya que pueden ser útiles en situaciones como las siguientes (Dukes, 1965; Vázquez Gómez, 1985; Bernia, 1981):

1) Cuando se requieren estudios idiográficos, es decir, estudios intensivos de individuos.
2) Cuando está justificado suponer que el proceso estudiado es ya general porque las diferencias individuales son mínimas.
3) Cuando sólo hay oportunidad de observar y estudiar un solo sujeto.
4) Si se pretende estudiar en profundidad un problema, el estudio de un solo sujeto puede ayudar a definir variables, precisar las cuestiones y sugerir modos de abordarlo.

A continuación se describen los distintos tipos de diseños que hemos mencionado: grupos no equivalentes, series temporales interrumpidas y sujeto único.

6.2 DISEÑOS DE GRUPOS NO EQUIVALENTES

En ocasiones, los diseños de un solo grupo pueden ser viables para el educador, siempre que se tomen las debidas precauciones y se utilice la información con carácter exploratorio y provisional, ya que la comparación se ha de efectuar con un grupo hipotético. Los diseños de dos grupos, en cambio, pueden aportar resultados más válidos y fiables, pues permiten la comparación entre dos grupos de datos reales.

DISEÑOS DE GRUPO ÚNICO

Estos diseños constan de un solo grupo al que se aplica un tratamiento con medidas que pueden ser pretest-postest (tabla 6.1) o sólo postest (tabla 6.2).

TABLA 6. 1 Diseño de grupo único pretest-postest

Grupos	*Sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n	No azar	$\bar{X}_1$	a_1	$\bar{X}_2$

El diseño de grupo único pretest-postest permite comparar las medidas antes y después del tratamiento. Sin embargo, el control tiene una doble deficiencia. Por un lado, al manipular la variable independiente sólo se aplica un nivel de dicha variable, y por otro, existen muchas variables extrañas que han podido sumarse a la acción de la independiente. En consecuencia, el diseño tiene una validez interna muy baja.

Cuando no se ha podido efectuar la medición antes de aplicar el tratamiento (tabla 6.2), la validez es todavía más deficiente que en el diseño anterior y hay que comparar el nivel obtenido después de aplicar la variable independiente con un nivel hipotético.

TABLA 6. 2. Diseño de grupo único sólo postest

Grupos	Sujetos	Asignación	Pretest	Tratamiento	Postest
1	n	No azar	-	a_1	$\bar{X}_1$

Sin embargo, a pesar de la poca validez de ambos diseños, en algunas situaciones educativas puede ser la metodología más viable, por ejemplo, investigaciones sobre los cambios de actitudes en el profesorado al implantar un determinado proyecto educativo o cambios en el rendimiento al implantar la informática en un centro en la asignatura de lengua.

Para ilustrar este diseño supongamos que un educador quiere constatar la eficacia del dictado para mejorar la ortografía de sus alumnos de primaria. Durante un curso emplea el dictado en días alternas y en junio aplica la prueba de ortografía de A. Galí, con el fin de comparar los resultados obtenidos en su clase con los que hipotéticamente deberían obtener, o con los resultados recordados o registrados de manera informal en otros grupos de años anteriores. Sin embargo, dado que el educador no dispone de una medida pretest y no aplicó ningún otro nivel de la variable independiente, la validez interna y externa es tan baja que no puede considerarse una investigación científica. El nivel obtenido en ortografía puede ser debido a una gran multiplicidad de variables, como la selección diferencial de los sujetos, la historia y la maduración. Si hubiera llevado a cabo un diseño pretest -postest debería tenerse en cuenta también la reactividad de la medida y la regresión a la media.

DISEÑOS DE DOS GRUPOS

Este diseño permite comparar la medida de la variable dependiente del grupo sometido a un nivel de la variable independiente con la medida obtenida por otro grupo que no ha recibido dicho nivel de la variable independiente. Si se utiliza un diseño pretest-postest (tabla 6.3), puede estimarse la equivalencia de los grupos sólo en la variable que se mida, pero no en las demás.

TABLA 6. 3. Diseño de dos grupos pretest-postest

Grupos	Sujetos	Asignación	Pretest	Tratamiento	Postest
1	n_1	No azar	$\bar{X}_1$	a_1	$\bar{X}_3$
2	n_2	No azar	$\bar{X}_2$	a_0	$\bar{X}_4$

En caso de utilizar un diseño sólo postest (tabla 6.4), las posibles diferencias iniciales entre los grupos todavía pueden contaminar más los resultados.

TABLA 6. 4 Diseño de dos grupos sólo postest

Grupos	*Sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n_1	No azar	-	a_1	$\bar{X}_1$
2	n_2	No azar	-	a_0	$\bar{X}_2$

6.3 DISEÑOS DE SERIES TEMPORALES INTERRUMPIDAS

Básicamente, estos diseños constan de las siguientes fases:

1) Medir la variable dependiente en uno o varios grupos durante un período suficiente para que las medidas muestren cierta estabilidad o regularidad. Los grupos pueden ser siempre los mismos o pueden ser grupos similares.
2) Aplicar la variable independiente.
3) Medir la variable dependiente nuevamente.

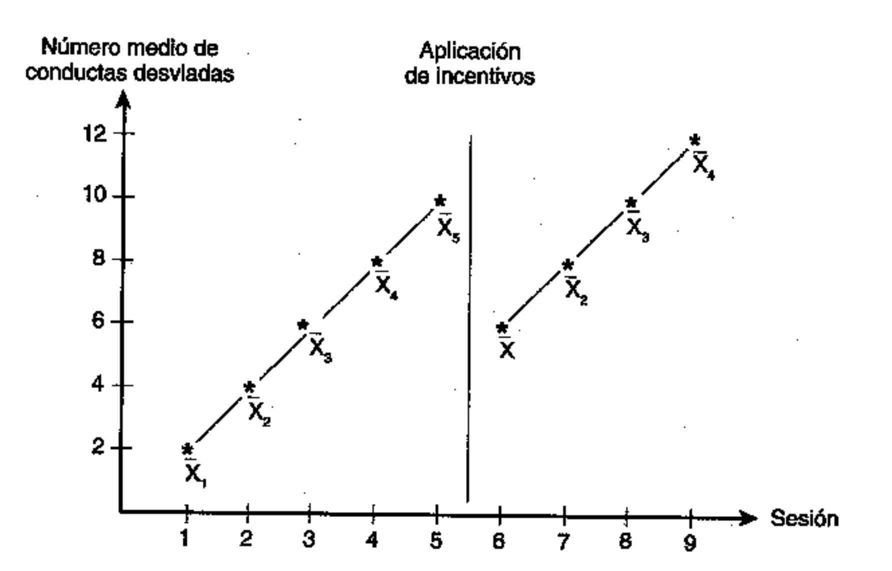
El efecto de la variable independiente puede estimarse analizando el grado de continuidad que presentan las medidas antes y después de aplicar la variable independiente.

Para ilustrar posibles efectos que se pueden presentar nos basaremos en un ejemplo. El número medio de conductas desviadas de un grupo de sujetos puede mostrar una discontinuidad a partir del momento en que se aplican determinados incentivos. Supongamos que se calculan cinco medias durante la fase anterior (1) a la aplicación de los incentivos y cuatro medias durante la fase posterior (2).

El análisis de los resultados del diseño puede efectuarse mediante el análisis visual de las representaciones gráficas, o bien con la técnica estadística ARIMA (AutoRegressive Integrated Moving Average), que permite describir un valor como una función lineal de datos anteriores y errores debidos al azar. En los paquetes estadísticos BMDP y SPSSX este análisis se denomina *prueba de Box-Jankins.*

Como puede observarse a través de la inspección visual del gráfico que se adjunta, existe una discontinuidad en la serie de medias que se han registrado.

Sesión	*Fas*	*Media conductas*
1	1	2
2	1	4
3	1	6
4	1	8
5	1	10
6	2	6
7	2	8
8	2	10
9	2	12



Si no se hubieran aplicado los incentivos, el nivel medio de conductas desviadas tendería a ser 12, 14, etc. Sin embargo, después de aplicarlos, el nivel medio de conductas desviadas desciende bruscamente a seis y sigue una tendencia similar a la que se registraba antes de aplicar los incentivos. El efecto se denomina *cambio de nivel*, sin que se haya visto afectada la inclinación o tendencia de la serie de medias. Sin embargo, también pudieran darse cambios que afectaran a la tendencia, de manera instantánea o permanente, o de forma retardada o pasajera.

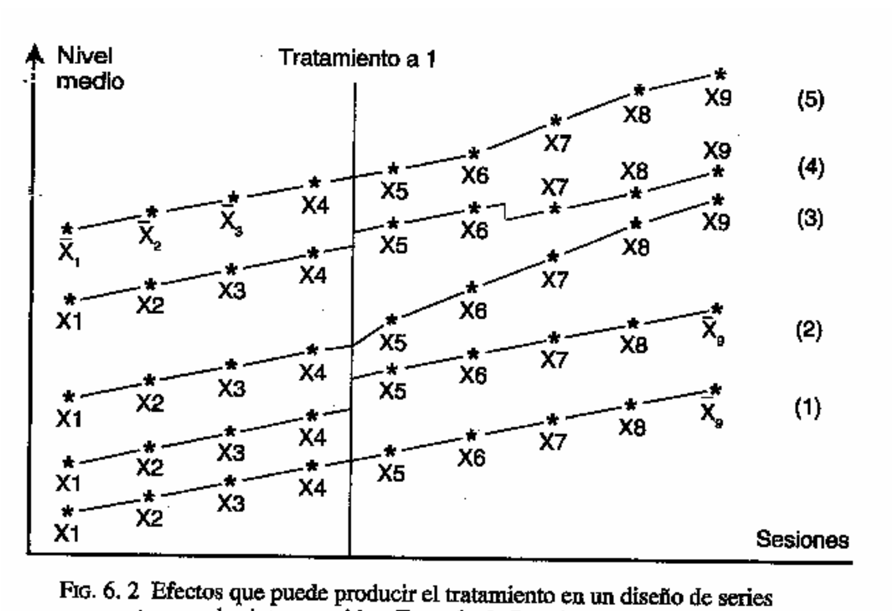


FIG. 6. 2 Efectos que puede producir el tratamiento en un diseño de series temporales interrumpidas. Tomado de Pereda (1987, 283)

3)

Los efectos que aparecen en la figura 6.2 pueden describirse del siguiente modo:

1) No se han producido efectos.
2) Cambio de nivel instantáneo y permanente o continuo.
3) Cambio de tendencia instantáneo y permanente o continuo.
4) Cambio de nivel instantáneo y pasajero.
5) Cambio retardado y pasajero en la tendencia o inclinación.

A continuación se exponen los principales diseños de series temporales interrumpidas que se utilizan con más frecuencia a nivel cuasiexperimental.

DISEÑOS SIMPLES

Se recoge información de un solo grupo de sujetos, efectuándose medidas antes del tratamiento y después de su aplicación. Algunas fuentes de invalidez importantes son la maduración, la historia y la reactividad de la medida. La representación simbólica del diseño es:

Grupos	*Número de sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n	No azar	$\bar{X}_1, \bar{X}_2, \bar{X}_3$	a_1	$\bar{X}_4, \bar{X}_5, \bar{X}_6$

DISEÑO DE DOS GRUPOS NO EQUIVALENTES

Este diseño incluye un grupo control en el que se toman las mismas medidas que en el grupo que recibe el tratamiento. No se puede asumir la equivalencia entre ambos grupos, pero es posible controlar algunas fuentes de invalidez como la madurez y la historia. Su representación simbólica es la siguiente:

Grupos	*Número de sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*
1	n_1	No azar	$\bar{X}_1, \bar{X}_2, \bar{X}_3$	a_0	$\bar{X}_4, \bar{X}_5, \bar{X}_6$
2	n_2	No azar	$\bar{X}_7, \bar{X}_8, \bar{X}_9$	a_1	$\bar{X}_{10}, \bar{X}_{11}, \bar{X}_{12}$

DISEÑO CON RETIRADA DE TRATAMIENTO

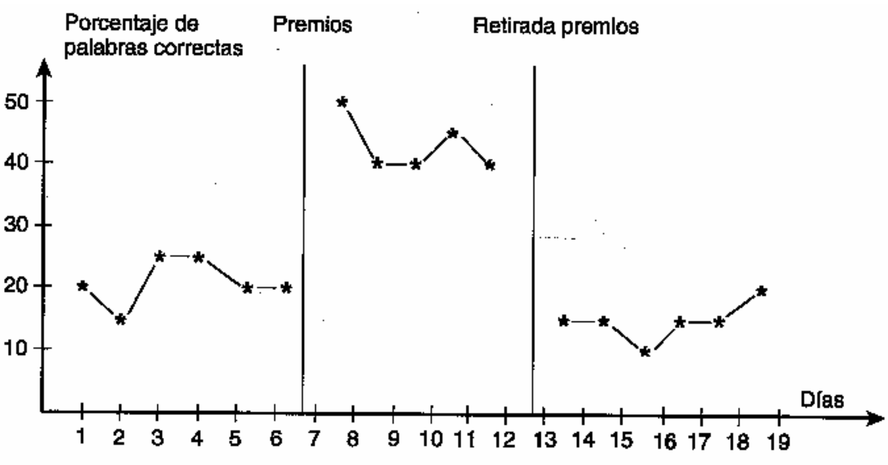
Con respecto a los diseños anteriores, el diseño con retirada del tratamiento tiene la particularidad de que dicho tratamiento se deja de aplicar a los sujetos en un

momento determinado y el investigador sigue tomando medidas de la variable dependiente. Al retirar el tratamiento los sujetos vuelven a la situación inicial. Esta modalidad permite un mayor control de las variables extrañas porque el investigador puede comparar dos series de medidas antes y después del tratamiento con una serie de medidas obtenidas durante la aplicación del tratamiento. Este diseño puede representarse del siguiente modo:

Grupos	*Sujetos*	*Asignación*	*Pretest*	*Tratamiento*	*Postest*	*Retirada*
1	n_1	No azar	$\bar{X}_1, \bar{X}_2, \bar{X}_3$	a_1	$\bar{X}_4, \bar{X}_5, \bar{X}_6$	$\bar{X}_7, \bar{X}_8, \bar{X}_9$

La serie de medidas X_4, *Xs'* X_6 permite estimar los efectos del tratamiento, mientras que las medidas X?, X_8, X_9 tienen la finalidad de constatar los efectos de la ausencia del tratamiento.

Un ejemplo de representación gráfica asociada a este diseño puede ser la siguiente:

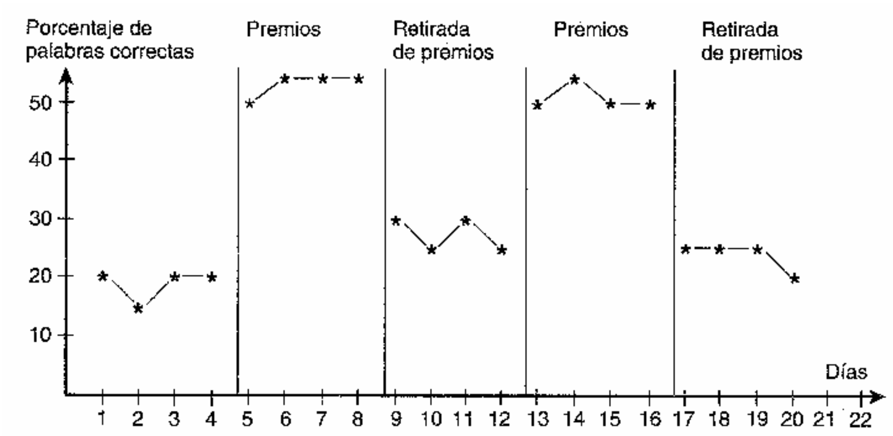


DISEÑO CON REPLICACIONES MÚLTIPLES

La estrategia utilizada en el diseño anterior puede repetirse sucesivamente en una misma investigación, dando lugar a la siguiente representación simbólica en la que se ha simbolizado el tratamiento por VI:

Gr.	*Asig.*	*Pretest*	*VI*	*Postest*	*Retirada*	*VI*	*Postest*
1	No Az	$\bar{X}_1, \bar{X}_2, \bar{X}_3$	a_1	$\bar{X}_4, \bar{X}_5, \bar{X}_6$	$\bar{X}_7, \bar{X}_8, \bar{X}_9$	a_1	$\bar{X}_{10}, \bar{X}_{11}, \bar{X}_{12}$

Este diseño ofrece más posibilidades de control que los anteriores, aunque es menos viable cuando los efectos de los tratamientos perduran después de su retirada. Un ejemplo gráfico de posibles resultados hipotéticos es el siguiente:



6.4. DISEÑOS DE SUJETO ÚNICO

Los diseños de sujeto único, también denominados *intrasujeto* o *de* $N = 1$, son similares a los diseños de series temporales interrumpidas, pero ahora no intervienen grupos de sujetos, sino un solo sujeto.

Numerosos trabajos y descubrimientos se basan en el estudio de un solo sujeto. Por ejemplo, los estudios de Fechner sobre umbrales sensoriales y diferencias mínimas perceptibles, de Wundt con la introspección, de Ebbinghaus sobre la memoria, de Kohler sobre la percepción distorsionada o de Pavlov sobre el condicionamiento clásico.

PROCEDIMIENTO GENERAL

La estrategia específica de estos diseños se basa en las siguientes consideraciones:

1) Describir las características del sujeto. Los resultados sólo serán generalizables a sujetos de características similares a las del sujeto estudiado. De ahí la necesidad de una descripción exhaustiva de la conducta durante la intervención educativa y antes de la misma tratando de buscar posibles variables que pudieran influir en los resultados.

2) Mediciones repetidas, efectuadas siempre en condiciones idénticas.

3) Elección de la línea base: ofrece un criterio de comparación del cambio producido en la variable dependiente por la variable independiente y cumplirá dos condiciones: longitud suficiente y estabilidad. Durante esta fase, denominada *línea base*,

período basal o *fase A*, se realizan una serie de observaciones o medidas antes de aplicar la variable independiente.

4) Aplicación de la variable independiente y observación o registro de la variable dependiente después de haberse iniciado el tratamiento o variable independiente. Dicha intervención educativa puede ser temporal (una sola aplicación puntual) o continua (prolongada). Este período se denomina *fase experimental, fase de tratamiento* o *bien fase B.*

5) Repetición sistemática: para aumentar la generalización de los resultados del experimento hay que repetido en varios sujetos similares, en sujetos de otras características, en otros contextos y con diferentes experimentadores.

A continuación se describen los diseños más frecuentes y posteriormente se ilustra el análisis visual de posibles resultados.

DISEÑO AB

Se realiza la observación y registro de la conducta o variable dependiente durante un período de tiempo determinado, y a continuación se introduce la variable independiente o tratamiento durante otro período, también con observación y registro de la variable dependiente. Este diseño no permite determinar con certeza la relación causal entre las variables independiente y dependiente, ya que las posibles diferencias entre las observaciones de ambos períodos (línea base y fase experimental) han podido ser provocadas también por otras variables. Algunas fuentes de invalidez interna pueden ser la maduración, la historia y la regresión estadística, ya que han podido provocar parte de las diferencias encontradas.

Los resultados de un diseño AB aparecen en la figura 6.3. Durante ocho días consecutivos un educador contabilizó el porcentaje de palabras bien pronunciadas por un niño esquizofrénico durante una sesión diaria de logopedia. A partir del noveno día, el educador inició un tratamiento basado en una actitud elogiante al concluir cada sesión.

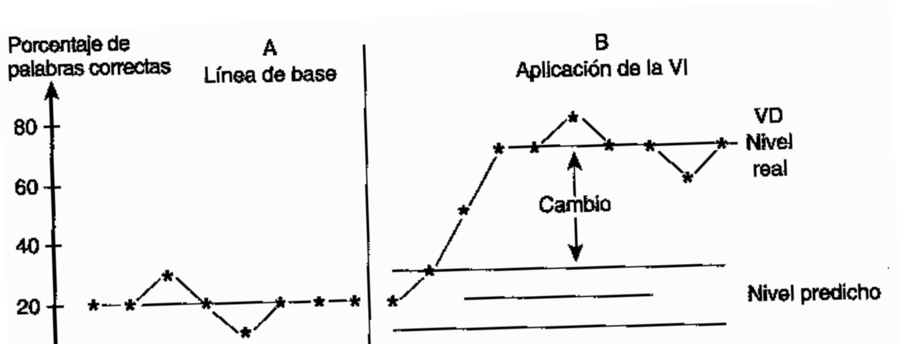


FIG. 6. 3. Representación gráfica de los datos obtenidos con un diseño AB

El nivel predicho indica la predicción de la conducta a partir de los datos precedentes en el caso de que no se hubiese producido la intervención educativa elogiante. Los registros realizados desde el día 9 al 18 representan el porcentaje de palabras correctas pronunciadas bajo la acción del tratamiento. La diferencia observada entre el nivel predicho y el nivel real da una estimación de la intensidad del cambio o efectividad del tratamiento.

DISEÑO DE LÍNEA DE BASE MÚLTIPLE

Como el diseño anterior, se compone de observaciones o registros en series temporales. Puede establecerse un paralelismo con los diseños de un solo grupo con pretest y postest, pero teniendo en cuenta que ahora el «grupo» está formado por un solo sujeto. Las fases son las siguientes:

1) Período Al: se registran observaciones que configuran la línea base de la conducta que va a ser la variable dependiente.
2) Período B: se introduce la variable independiente y se registran los valores de la variable dependiente.
3) Período A_2: se retira la variable independiente y se siguen registrando los valores de la variable dependiente.

Imaginemos que en el ejemplo anterior se retira la actitud elogiante a partir del día 17. Los registros obtenidos aparecen en la figura 6.4.

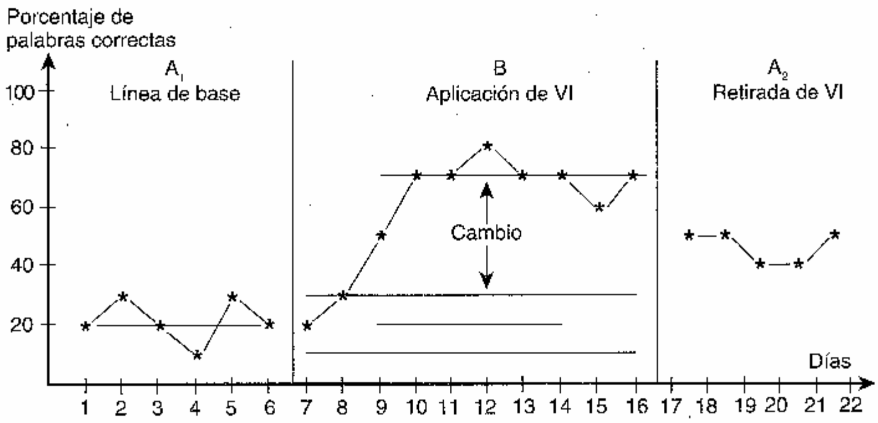


FIG. 6. 4. Representación gráfica de los datos obtenidos con un diseño ABA

Si hay diferencias entre los valores de la fase Al y los correspondientes de la fase B, y durante el período A_2 se aprecia un retorno a los niveles del período Al' podemos atribuir las diferencias al efecto de la variable independiente. Sin embargo, existe la posibilidad de que la variable dependiente siga una serie de cambios cíclicos y que los

períodos elegidos coincidan con esos ciclos. De ahí la importancia de que la línea base y la fase de retirada de la variable independiente se prolonguen lo necesario hasta que se estabilice la variable dependiente.

DISEÑO DE LÍNEA BASE MÚLTIPLE

Constituyen una extensión del diseño AB y son una alternativa a los diseños ABA cuando no es adecuado suprimir o retirar la variable independiente. Sus características son las siguientes:

1) Emplean más de una variable dependiente (VD).
2) Hay que observar o registrar las conductas (VD) durante un período A, obteniéndose las líneas base correspondientes a cada conducta (VD).
3) Se aplica la variable independiente (VI) a una de las conductas (VD¡), posteriormente se aplica a la segunda conducta (VD), y luego a la tercera (VD$_3$). Es decir, la misma VI se aplica a cada VD, pero la intervención se realiza en períodos diferentes.

Supongamos que durante determinados períodos de las clases de música la profesora considera que dibujar, hablar y leer son tres conductas de un sujeto que no son deseables (fig. 6.5). Para reducidas o eliminarlas introduce un programa consistente en dedicar tiempo de recreo a ensayar personalmente con el sujeto, ayudándose de un instrumento musical. El tiempo del ensayo era sucesivamente proporcional a la ausencia de las distintas conductas implicadas. Primero se tenía en cuenta el porcentaje de tiempo en que no se dibujaba, luego se hacía lo mismo con respecto al habla y con la lectura. Las distintas conductas se utilizan como controles; por ello es conveniente que sean conductas independientes, es decir, no correlacionadas.

Como puede verse en las representaciones gráficas (fig. 6.5), los ensayos durante los

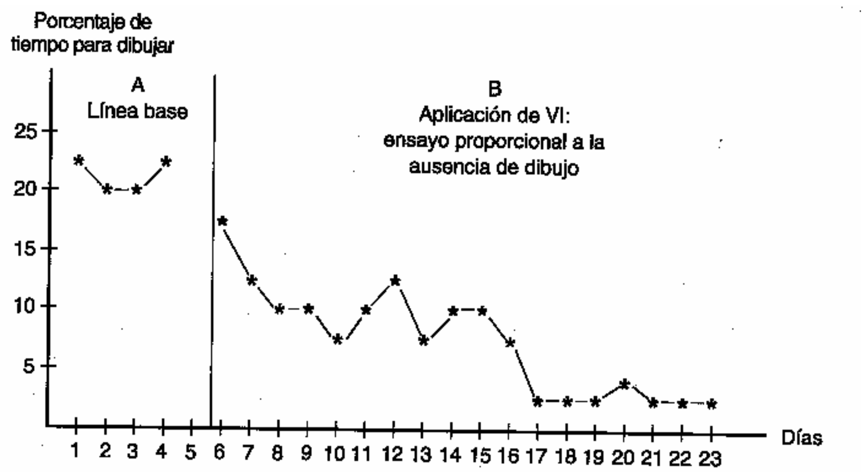


FIG. 6. 5 (continúa)

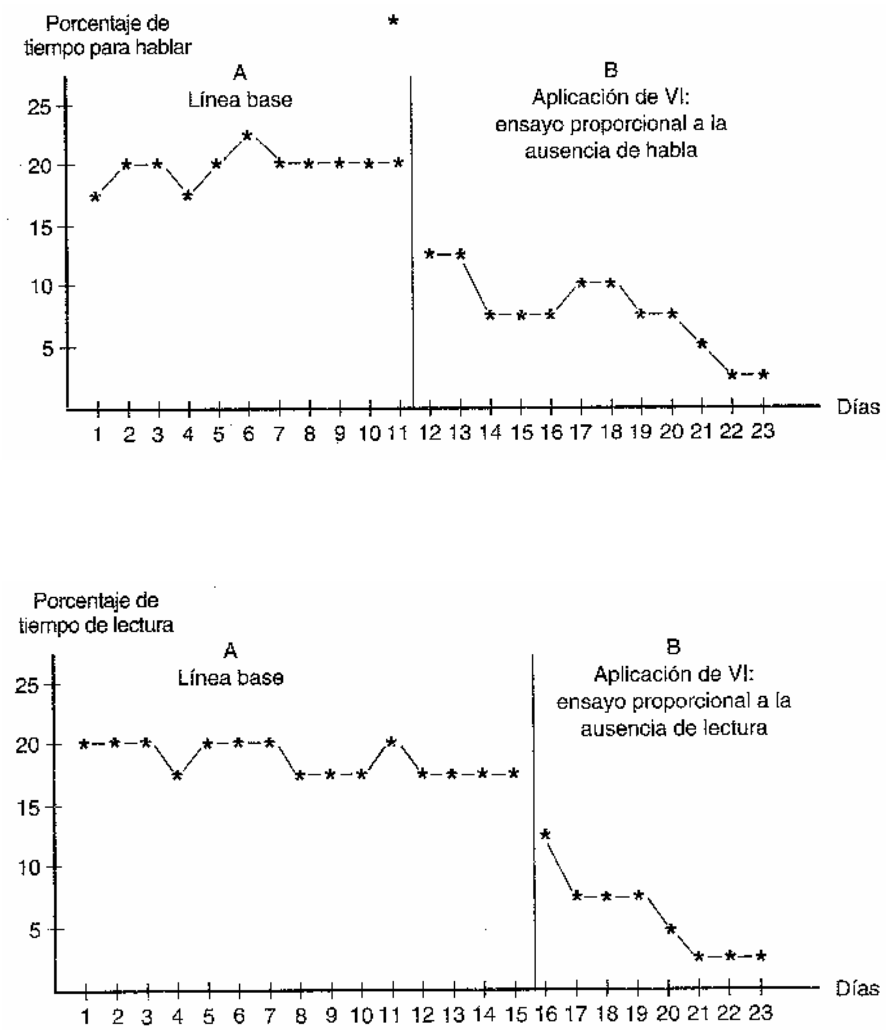


FIG. 6. 5 Representación gráfica de los datos de un diseño de línea base múltiple

recreos van seguidos de una disminución en la conducta correspondiente pero no en las demás, lo que permite atribuir el cambio ala variable independiente con mayor seguridad.

El esquema puede utilizarse para estudiar una misma conducta, pero empleando sujetos diferentes aunque de características similares. También puede estudiarse una sola conducta en situaciones distintas.

ASNÁLIS VISUAL DE LOS RESULTADOS

Para medir la variable dependiente hay que tener en cuenta que puede tratarse de conductas manifiestas o directamente observables (en ocasiones, indicadores de constructos psicológicos subyacentes), de conductas (respuestas) verbales, o incluso

de respuestas fisiológicas. En dichas conductas podemos contabilizar la frecuencia de su aparición (por cada sesión o unidad de tiempo), el tiempo de latencia (tiempo que tarda en ser emitida una conducta), la duración de la respuesta y la intensidad de la respuesta.

Un ejemplo de recogida de datos, con intervención a partir de la sesión 5 y retirada a partir de la 10, sería el que se ha representado en la figura 6.6.

El análisis de los resultados puede llevarse a cabo a través de un análisis visual

Sesión	*Fase*	*Tiempo estudio*
1	1	20
2	1	20
3	1	30
4	1	20
5	2	70
6	2	60
7	2	70
8	2	70
9	2	80
10	3	50
11	3	50
12	3	50
13	3	40

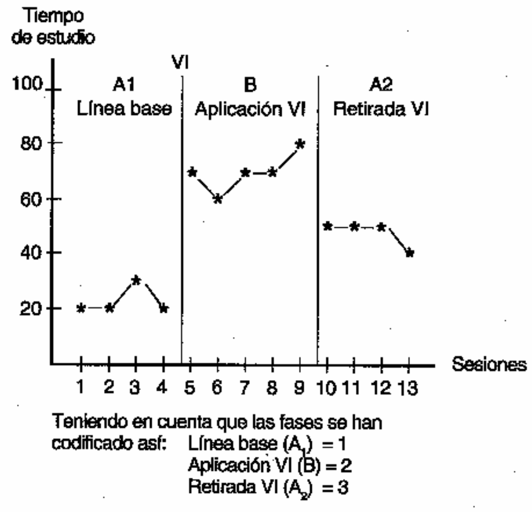


FIG. 6. 6 Matríz de datos de un diseño ABA y representación gráfica

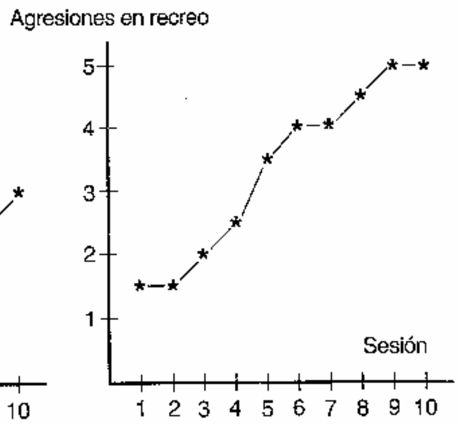
y/o estadístico (modelo ARIMA). Para este análisis estadístico remitimos a obras como la de Arnau (1984).

Por lo que se refiere a la inspección visual de las representaciones gráficas, conviene analizar cambios dentro de las fases y entre fases. Estos últimos pueden ser de nivel con línea base estacionaria o no estacionaria, y cambios de tendencia, pero no de nivel con línea base estacionaria o no estacionaria. Algunos ejemplos son los siguientes.

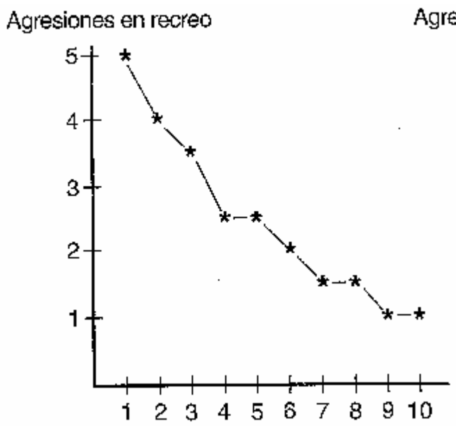
Cambios dentro de las fases



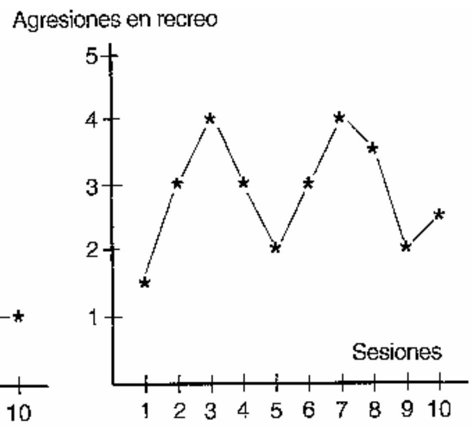
Sin orientación determinada. Las variaciones son mínimas y atribuibles al azar. Línea estable o estacionaria. Si se trata de una línea de base es apropiada para poner de manifiesto el efecto de una variable independiente.



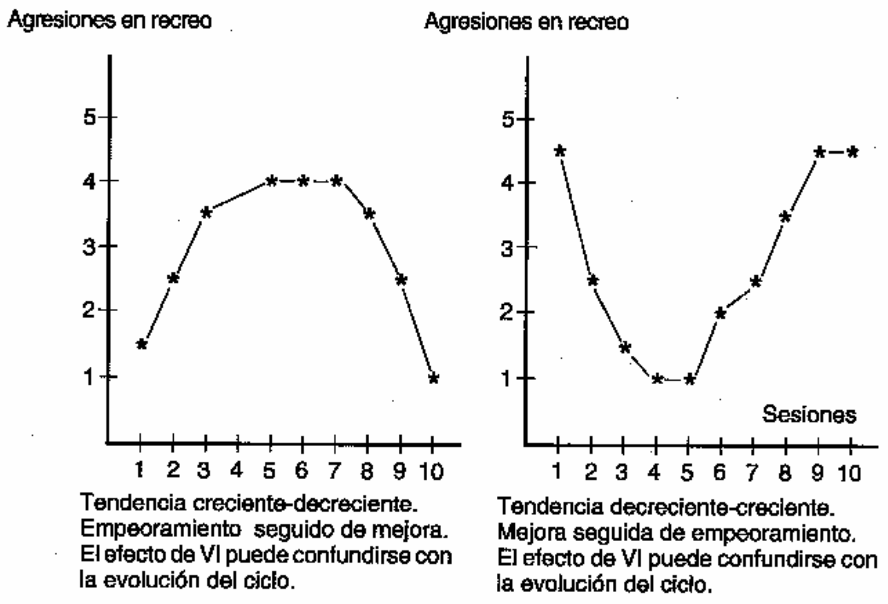
Tendencia a un empeoramiento progresivo. Línea deteriorante. En caso de ser una línea base es una situación ideal para introducir una variable independiente que tenga un efecto contrario.



Tendencia a la mejora. Dificultad para poner de manifiesto el efecto de la variable independiente.



Variabilidad cíclica: grandes cambios y de signo contrario. Dificultad para poner de manifiesto el efecto de VI.

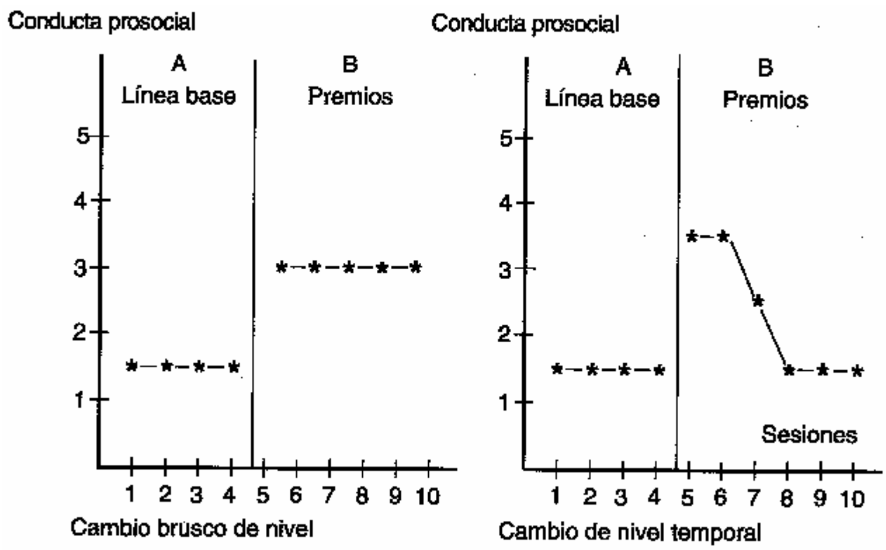


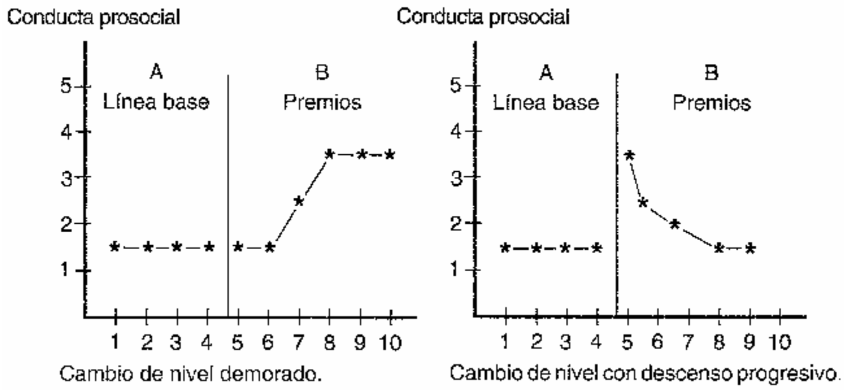
Cuando las observaciones no siguen una tendencia determinada es necesario proseguir con el registro hasta conseguir su completa estabilidad.

Cambios entre fases

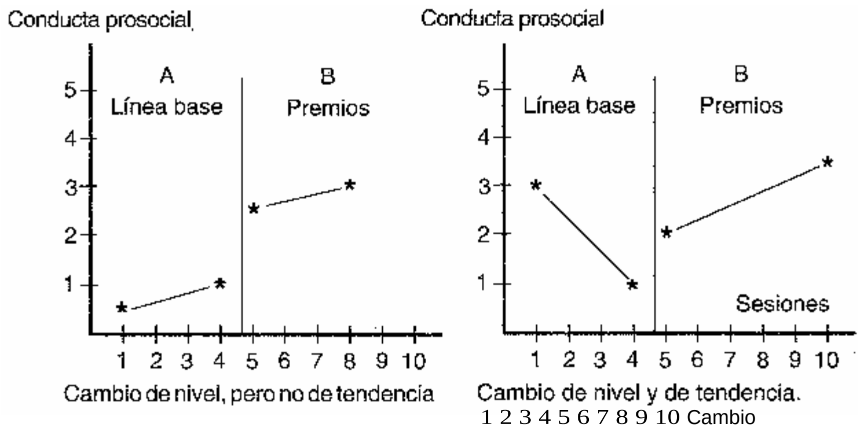
Pueden ocurrir:

a) *Cambios de nivel con línea base estable* o *estacionaria*



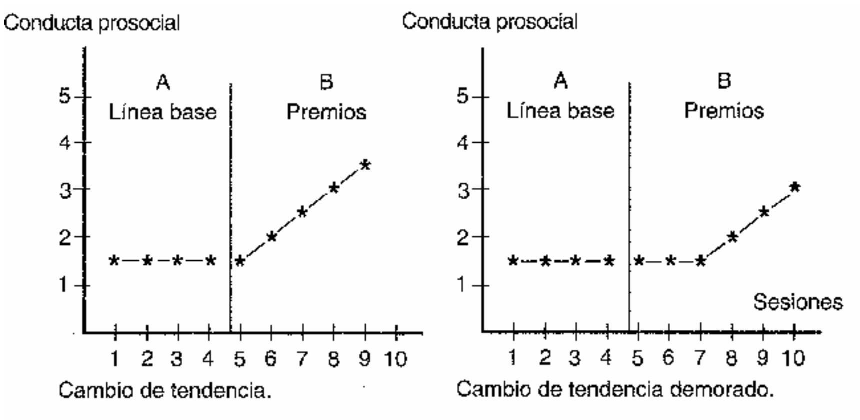


b) *Cambios de nivel con línea base no estacionaria (sigue una tendencia)*

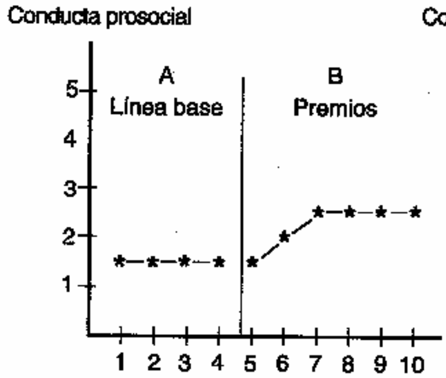


1 2 3 4 5 6 7 8 9 10 Cambio de tendencia demorado.

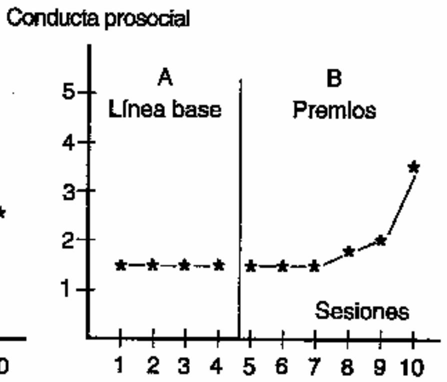
e) *Cambios de tendencia, pero no de nivel, con línea base estacionaria*



4)

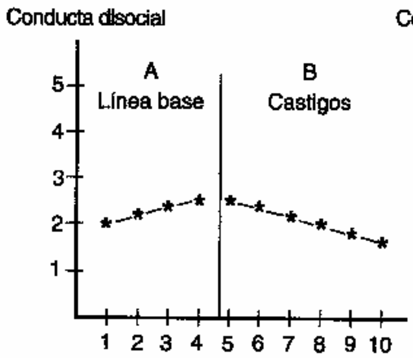


Cambio temporal de tendencia.

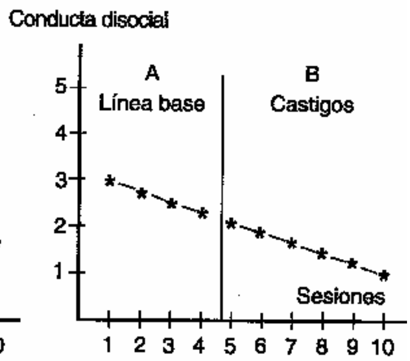


Cambio acelerado de tendencia.

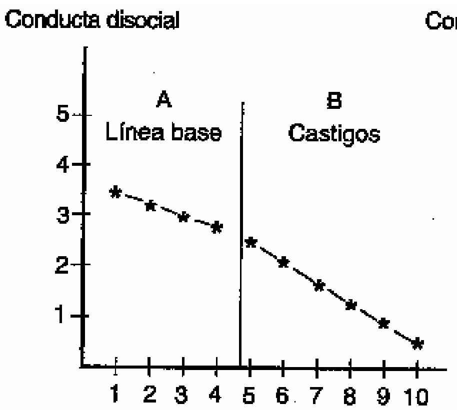
d) *Cambios de tendencia, pero no de nivel, con línea base no estacionaria*



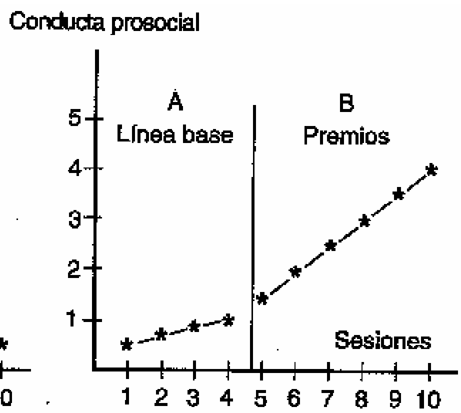
Cambio de dirección. Intervención efectiva.



Línia conductual sin cambio de inclinación. Dificultad para estimar la efectividad de la intervención educativa.



Mayor inflexión. La intervención acelera el proceso ya iniciado.



Igual que la anterior. La intervención acelera el proceso ya iniciado.

6.5 POSIBILIDADES Y LÍMITES

En general, el método cuasiexperimental está más indicado cuando la investigación se desarrolla en escenarios educativos naturales y se acepta la carencia de un control experimental completo. Esta deficiencia de control puede compensarse con la repetición de investigaciones (réplicas) o con observaciones y mediciones múltiples, con el fin de minimizar o incluso eliminar el mayor número posible de fuentes de invalidez interna.

a) *Posibilidades.* A través de la metodología cuasiexperimental también es posible explorar posibles relaciones causa-efecto, a pesar de que este objetivo es más propio del enfoque estrictamente experimental, caracterizado por el máximo control posible de las variables extrañas.

Esta metodología es muy importante en el ámbito educativo. Ofrece muchas ventajas por su proximidad a la realidad educativa, donde es frecuente que no se pueda realizar una investigación experimental al no ser viable alterar la estructura o configuración de grupos ya formados, con lo que es difícil poder aleatorizar los sujetos e incluso los tratamientos. Se aplica en las aulas, en los centros educativos y en otros contextos educativos donde las variables actúan en su propio medio natural. Esta dimensión real confiere a las variables la posibilidad de que ejerzan una influencia más potente que en un medio más artificial como el de un experimento. De ahí que este enfoque sea adecuado para estudiar las influencias sociales complejas, los procesos y cambios educativos en situaciones reales. Por otro lado, también permite poner a prueba las teorías y la solución de problemas prácticos.

b) *Límites.* Como contrapartida, junto a la falta de control y precisión suelen aparecer objeciones ante la manipulación de variables que pueden afectar al proceso educativo. Uno de los inconvenientes del enfoque cuasiexperimental es que muchas relaciones causa-efecto han podido producirse antes de la actuación del investigador y, por tanto, quedan fuera de su alcance. Aunque siempre se manipula la variable independiente, al no provocar directamente todo el fenómeno, la acción de algunas variables escapa a su control. Por ejemplo, cuando no existe la posibilidad de aleatorizar los grupos de sujetos, tenemos que realizar la intervención educativa sobre grupos naturales. En consecuencia, existen muchas variables potenciales que pueden influir de forma diferenciada en cada grupo y que escapan a nuestro control porque su influencia ya se ha podido producir con antelación.

Al tratarse de un enfoque cuasiexperimental no podemos saber con tanta certeza si han sido los niveles de la variable independiente manipulada los únicos responsables de los cambios observados al medir la variable dependiente en los distintos grupos. Haría falta recurrir a criterios de orden lógico. Pero a pesar de ello, la comparación de los grupos de datos permite llegar a conclusiones sobre la eficacia relativa de cada nivel de la variable independiente, siempre con las reservas mencionadas.

CAPÍTULO 7

Metodología no experimental

En el marco de la orientación empírico-analítica hemos contemplado un continuo metodológico que se extiende desde la metodología experimental hasta la ex-post - facto o no experimental. Ambas vías metodológicas presentaban aspectos bien diferenciados, tal como queda sintetizado en la tabla siguiente.

Metodología experimental	*Metodología no experimental*
Provocamos (manipulamos) los efectos	Los efectos ya se han producido.
Modificamos VI y observamos cambios en VD.	No se modifica, sólo seleccionar y observar.
Orientación hacia el futuro.	Orientación hacia el pasado.
Aleatorización de grupos.	Grupos naturales ya formados.

Cuando el investigador no dispone de la información necesaria para solucionar el problema planteado, puede crear o provocar el fenómeno como ya se ha descrito en los métodos experimental y cuasiexperimental, pero también puede buscar un contexto o situación donde obtener los datos necesarios porque el fenómeno ya se haya producido. Tendrá que acudir a la mencionada situación, recogerá los datos y los analizará, pero no modificará ni se provocará dicha situación, pues mantendrá una actitud pasiva. Los métodos no experimentales o ex-post - facto se limitan a describir una situación que ya viene dada al investigador, aunque éste pueda seleccionar valores para estimar relaciones entre las variables.

En sentido amplio, Kerlinger (1985, 268) conceptualiza la investigación ex-post-facto como una búsqueda sistemática empírica en la cual el científico no tiene control directo sobre las variables independientes porque ya acontecieron sus manifestaciones o por ser intrínsecamente no manipulables. Se hacen inferencias sobre las relaciones, sin intervención directa, a partir de la variación común de las variables independientes y dependientes.

7.1 TIPOS DE MÉTODOS

Bajo la denominación genérica de metodología *ex-post-facto o no experimental,* tal como se ha conceptualizado aquí, pueden incluirse diversos métodos, como los siguientes:

a) *Método comparativo-causal.* Esta modalidad se utiliza cuando el investigador intenta explicar relaciones de causalidad comparando grupos de datos, pero la variable que el investigador estudia como posible causa de los cambios observados en el criterio no es manipulable y sólo admite un nivel de selección. Sin embargo, también puede ocurrir que, siendo una variable manipulab1e, no pueda provocarse el fenómeno por razones de carácter ético, economía de tiempo y distorsión de la situación educativa. En estos casos está indicada la replicación continuada de las investigaciones.

b) *Métodos descriptivos.* Exploran relaciones y, para ello, tratan de asociar y comparar grupos de datos. Ejemplos de métodos descriptivos son los estudios de *desarrollo o evolutivos, de encuesta, de casos y observacionales.* Recogen y analizan información con fines exploratorios y pueden constituir una aportación previa para orientar futuros estudios correlacionales, predictivos o de corte experimental, que tratarán de contrastar las hipótesis generadas por aquéllos. Debe tenerse en cuenta que el estudio de *casos*, por su orientación más *idiográfica,* será abordado en la perspectiva *humanístico-interpretativa.*

c) *Métodos correlacionales.* Están indicados cuando el investigador busca el grado de relación entre variables e informa también sobre la variabilidad de una variable que queda explicada por otra variable. Este método permite explorar hasta qué punto las variaciones observadas entre las variables dependen unas de otras. Si la magnitud de la relación es suficiente puede derivarse un estudio predictivo. A partir de la matriz de correlaciones puede generarse un análisis factorial con el fin de explicar un número de variables mediante un número más reducido de variables subyacentes.

En los próximos apartados se describen los tres tipos de métodos *no experimentales* que se han mencionado: comparativo-causal, descriptivo y correlacional.

7.2. MÉTODO COMPARATIVO-CAUSAL

Muchos procesos, prácticas, programas y resultados educativos han sido analizados con este método, también denominado *explicativo causal* o *selectivo-comparativo.* Suelen abordarse interrogantes de índole educativa como pueden ser:

- ¿Por qué algunos niños leen mejor que otros?
- ¿Por qué algunos alumnos llegan a ser delincuentes y otros no?
- ¿Qué método didáctico influye más en el rendimiento?
- ¿ Qué efecto tiene la asistencia de niños de tres a seis años al aula de preescolar en la madurez lectora?
- ¿Existe relación entre el hecho de ejercer una profesión y el autoconcepto?

En el ámbito educativo no sólo interesa saber cómo es un fenómeno, sino también de qué manera y por qué ocurre. En consecuencia, se comparan las semejanzas y

diferencias que existen entre los fenómenos para descubrir los factores (variables o condiciones) que parecen acompañar o contribuir a la aparición de ciertos hechos y situaciones en su propio contexto natural.

En general, si se tiene en cuenta cuándo ocurrió la relación entre las variables y la posibilidad de manipular la variable independiente, pueden darse dos situaciones de investigación que pueden analizarse mediante el método selectivo-comparativo:

1) Cuando la posible influencia de la variable independiente sobre la variable dependiente ya se ha producido al comenzar la investigación. Por ejemplo, si queremos medir el efecto de la educación infantil en el desarrollo de la madurez lectora, la variable independiente es susceptible de manipulación y sería posible realizar un experimento. Sin embargo, además de que sería inviable asignar cada niño a una de las dos categorías de la variable independiente, tendríamos que prolongar la investigación durante un curso escolar.

La investigación es mucho más breve si pretendemos estudiar la cuestión con niños que actualmente ingresan en primaria. Es evidente que la influencia de los cursos previos de educación infantil ya se ha producido con anterioridad y sólo podemos seleccionar dos muestras de niños (una que haya recibido educación infantil y otra no) y comparar actualmente su madurez lectora. Muchos problemas de investigación pedagógica se ajustan a este caso.

2) Cuando se desea investigar la influencia de variables no susceptibles de manipulación como el sexo, el ambiente familiar, la motivación, la inteligencia, los hábitos de lectura de los padres y de otras personas, es imposible asignar los sujetos aleatoriamente a diferentes categorías de dichas variables. Se pueden elegir al azar sujetos que posean determinados valores de la variable independiente, pero la relación de cada sujeto con el valor de la variable independiente no es al azar porque ya se había producido cuando se seleccionan los sujetos, es decir, sus manifestaciones ya acontecieron.

Si la variable independiente es manipulable, puede aplicarse la metodología experimental o cuasiexperimental, pero cuando las variables no son manipulables, o no interesa manipularlas, hay que utilizar el método comparativo-causal o selectivo-comparativo. La mayoría de los estudios en psicología y pedagogía diferencial son de este tipo.

El investigador logra la variación que desea, no por manipulación directa de la variable independiente, sino seleccionando sujetos en los que la variable independiente está presente o ausente, es más intensa o menos intensa, fuerte o débil, etc., y compara los valores de la variable dependiente en cada grupo. Mientras que en el método experimental provocamos la ocurrencia de la variable dependiente y observamos su repercusión en la variable dependiente, en el método comparativo-causal observamos la ocurrencia de la variable dependiente y comprobamos si se ha dado anteriormente la ocurrencia de la variable independiente. Es decir, en un experimento el investigador puede suponer que si se somete a los sujetos a la variable independiente, se observará el resultado en la variable dependiente. En consecuencia, decide manipular la variable independiente; para ello, expone a un grupo experimental a un nivelo categoría de variable independiente, y compara los resultados con respecto a un grupo control.

En un estudio comparativo-causal, el proceso es inverso: parte de la observación o medición de la variable dependiente cuando ya ha recibido la posible influencia de

la variable independiente. Para ello selecciona dos o más grupos según el número de categorías o niveles de la variable independiente, y trata de hallar, entre las múltiples causas posibles o posibles variables independientes, que se relacionan con la variable dependiente o contribuyen a determinar el cambio observado. El investigador analiza una situación vital en la cual los sujetos han experimentado el fenómeno que quiere investigar.

En el ejemplo de la madurez lectora, el diseño podría ser:

Grupos	*n*	*Ed. Infantil*	*Lectura*
1º A	35	Sí	20
1º B	30	No	15

Es evidente que el método experimental aplicado al ámbito educativo no proporciona una garantía absoluta sobre el grado de control, pero en el comparativo-causal la incertidumbre es aún mayor. Así, en el ejemplo planteado, además de diferir los sujetos en haber recibido o no la educación infantil, también pueden diferir en otras variables que potencialmente hubieran podido provocar el cambio observado en la madurez lectora.

Por otra parte, no siempre es fácil establecer la dirección de la relación de causalidad. Por ejemplo, supongamos que se pretende analizar si el hecho de desempeñar una profesión está relacionado con el autoconcepto, y al comparar el autoconcepto de los sujetos que trabajan y los que no trabajan comprobamos que aquéllos presentan significativamente un mayor nivel en el mencionado rasgo psicológico. Sin embargo, cabe plantearse: ¿el hecho de ejercer una profesión ha provocado el cambio en el autoconcepto?; o por el contrario, ¿el hecho de poseer un mayor autoconcepto ha provocado que los sujetos hayan conseguido una profesión?, o bien, ¿ha sido el mayor nivel de aptitudes o capacidades personales el que determina a la vez un mayor autoconcepto y la posibilidad de encontrar o forjar una profesión?; es decir, existe la posibilidad de que tanto el autoconcepto como el ejercer una profesión dependan a la vez de las aptitudes o capacidades personales y por ello aparecían relacionados.

Dado que el método comparativo-causal es muy utilizado en educación, conviene disminuir en la medida de lo posible estas ambigüedades. Además de realizar réplicas o repeticiones de la misma investigación pueden tenerse en cuenta una serie de condiciones necesarias para poder inferir relaciones de causalidad (Ary y otros, 1987, 286). Así, ha de existir una relación estadística entre las variables implicadas y una determinada secuencia temporal. Además, conviene examinar la independencia en relación con otras variables y la posible existencia de hipótesis alternativas, como pueden ser la causa común y la causalidad inversa. A continuación se clarifica cada una de estas condiciones.

a) *Relación estadística*

La relación ha de existir entre dos o más variables, constatada a través de un coeficiente de correlación, o bien comparando las medias de la variable dependiente

en función de las categorías de la supuesta variable independiente. Es una condición necesaria, pero no suficiente para poder inferir la existencia de una relación de causalidad.

b) *Secuencia temporal*

La variable independiente precede a la dependiente en el tiempo. En el caso de que el cambio en la variable dependiente ocurra antes de que la presunta variable independiente estuviera presente, habrá que concluir que dicha variable no podrá ser la causa de los cambios observados en la variable dependiente. Se toman decisiones acerca de la relación temporal entre ambas sobre una base lógica o como resultado de mediciones que muestren que los grupos no diferían en la variable dependiente antes de su exposición a la variable independiente.

c) *Independencia de otras variables*

Para que exista relación de causalidad entre la variable independiente y la variable dependiente ha de ocurrir que la variable dependiente no dependa de otras variables. Tendremos que examinar si, además de la variable independiente, otras variables causan las diferencias detectadas en la variable dependiente. Para verificar esta posibilidad se introducen otras variables dentro del análisis y se observa en qué forma afectan a la relación entre la variable independiente y la variable dependiente. Quizá se descubra que tal relación perdura aún si se introducen otras variables. En este caso, si se ha recurrido a las variables más relevantes, se dispone de pruebas que apoyan una inferencia de causalidad. Por otra parte, tal vez se descubra que la presencia de las otras variables puede cambiar la relación entre las variables independiente y dependiente o incluso eliminarla. De ser así, se saca la conclusión de que o la variable independiente no influye en la variable dependiente, o la relación entre ambas es casual o aparente (espuria) porque viene determinada por una tercera variable.

d) *Hipótesis alternativas*

Pueden plantearse con respecto a la variable independiente o con respecto a la variable dependiente. Esto exige un cuidadoso análisis del problema y de las posibles causas relacionadas con un efecto y de los posibles efectos relacionados con una causa, así como de las interacciones entre distintas variables. Estas hipótesis han de formularse teniendo en cuenta la posibilidad de que exista:

1) Una causa común. Las variables independiente y dependiente pueden aparecer relacionadas sólo porque ambas sean efectos de una tercera variable. Por ejemplo, en el caso de la profesión y el autoconcepto, la aptitud o capacidad personal sería una causa común si pudiera provocar a la vez que los sujetos consiguieran una profesión y mejoraran su auto concepto, lo que origina una relación entre las dos últimas variables. En un estudio comparativo-causal, el investigador siempre ha de sospechar la posibilidad de que una causa o causas comunes hayan originado la relación observada.

2) Causalidad inversa. En lugar de afirmar que la variable independiente es causa de la variable dependiente, quizá ocurra lo contrario. Las investigaciones sobre prácticas educativas han revelado que los niños que son castigados con frecuencia muestran una conducta más agresiva. Por ejemplo, algunos datos ficticios que miden el nivel de agresividad según el número de agresiones semanales a compañeros, podrían ser los siguientes:

Frecuencia de castigos paternos	*Nivel de agresividad*
Alta	10
Media	4
Baja	2

Además de la hipótesis de que el castigo provoca mayor agresividad, hay que tener en cuenta la hipótesis inversa: el hecho de que los niños sean más agresivos provoca que los padres les castiguen más.

Consideraciones similares deberían hacerse al relacionar el alcoholismo y el rendimiento en la universidad. Los siguientes datos ficticios parecen indicar que un grado alto de alcoholemia provoca un menor rendimiento:

Grado de alcoholismo	*Rendimiento medio*
Alto	4
Medio	6
Bajo	8

Pero también hay que tener en cuenta la hipótesis inversa de que sea el fracaso universitario el que motive un mayor consumo.

3) Otras variables independientes. Aparte de la variable independiente que estudiamos puede haber otras variables independientes que ocasionen el efecto observado en la variable dependiente. Hay que enumerar todas las posibles variables independientes alternativas. Se procurará mantener todas constantes menos una para constatar si se relaciona con la variable dependiente. Si logramos eliminar las variables independientes alternativas demostrando que no se relacionan con la variable dependiente, obtendremos apoyo para la hipótesis original de la relación de causalidad entre la variable independiente y la variable dependiente.

Si estamos interesados en explorar la posible incidencia de la clase social en el rendimiento de los alumnos, tendremos que seleccionar sujetos que pertenezcan a las distintas clases sociales, con el fin de comparar su rendimiento respectivo. Sin

embargo al seleccionar los sujetos según la clase social también han podido quedar seleccionados según otras variables extrañas o ajenas a la investigación, pero que han podido influir en la variable dependiente, como pudiera ser la inteligencia. Así, podría ocurrir con los siguientes datos ficticios, donde el posible efecto de la clase social puede quedar contaminado por la inteligencia:

Clase social	*Inteligencia media*	*Rendimiento medio*
Alta	110	7
Media	100	5
Baja	95	3

7.3. MÉTODOS DESCRIPTNOS

El método descriptivo se ha concebido de diversas maneras. Así, para Bartolomé (1984) el método descriptivo incluye el método observacional y el método exploratorio; en cambio, para otros autores (Mouly, 1978; Fox, 1981) el método exploratorio es el que engloba a los descriptivos. Según Travers (1971) y Van Dalen y Meyer (1980), el método descriptivo se identifica con el observacional. Otros autores plantean una tercera alternativa que denominan *survey* o *método exploratorio* (Mouly, 1978, y Fox, 1981).

En general, los distintos autores coinciden en que este método pretende describir un fenómeno dado, analizando su estructura y explorando las asociaciones relativamente estables de las características que lo definen, sobre la base de una observación sistemática del mismo, una vez producido (Echeverría, 1983). Es un método inductivo, siendo uno de sus objetivos describir hipótesis (Bartolomé, 1984).

La denominación de *método descriptivo,* en sentido amplio, es la más aceptada en nuestra comunidad científica (Bartolomé, 1984, y Orden Hoz, 1985) y en ella hemos incluido los estudios de desarrollo, que tienen como principal objetivo conocer los cambios que se producen en los sujetos con el transcurso del tiempo; las distintas matizaciones del método de encuesta (Mouly, 1978; Van Dalen y Meyer, 1980, y Bartolomé, 1984) orientadas a la descripción de una situación dada; el estudio de casos, más centrado en describir y analizar detalladamente unidades o entidades educativas únicas; y el método observacional, caracterizado porque la información es recogida de forma directa de los sujetos observados y no mediante sus respuestas.,

Al abordar la parte dedicada a la recogida de datos, nos referiremos nuevamente a la encuesta y a la observación consideradas como técnicas. Sin embargo, se incluyen aquí porque en el contexto de algunas investigaciones ambas técnicas adquieren tal relevancia que llegan a constituir un método con entidad propia. Por otro lado, como ya se había mencionado el estudio de casos, por su orientación predominantemente idiográfica, se tratará en el contexto de la perspectiva humanístico- interpretativa.

ESTUDIOS DE DESARROLLO

Este tipo de estudios describe la evolución de las variables durante un período de tiempo determinado. Así, un profesor puede describir las etapas de evolución del lenguaje, del dibujo, de la inteligencia o del juego en niños de primaria.

En general, los estudios de desarrollo o evolutivos se centran en diferencias ligadas a la edad. Así ocurre con muchas investigaciones en pedagogía diferencial, o en psicología evolutiva, como son las realizadas por J. Piaget sobre el pensamiento infantil, que han pasado a ser clásicas en la corriente cognitivo-evolutiva.

Los estudios de desarrollo pueden presentar las modalidades longitudinal, transversal y de cohortes.

a) *Estudios longitudinales*

En un estudio longitudinal se analizan características de los mismos individuos en distintos momentos o niveles de edad (E_j E_2, E_3, ... E), mediante observaciones repetidas ($0_{j,}$ 0_2, 0_3, ... 0). Supongamos que deseamos estudiar la evolución de la inteligencia con el test factor «g» de Cattell entre los 10 Y 15 años, podemos elaborar el siguiente diseño ilustrado con datos ficticios (tabla 7.1).

TABLA 7.1 Datos de un diseño longitudinal

Años	1985	1986	1987	1988	1989	1990
Edad	10	11	12	13	14	15
X en inteligencia	21	23	24	26	29	30
Nº de sujetos	40					

b) *Estudios transversales*

En un mismo momento se estudian distintos períodos evolutivos. Se comparan diferentes grupos de edad (G_j, G_2, G_3, ... G) observados en un único momento (al)' En el mismo ejemplo de la evolución de la inteligencia, los datos podrían ser los que se registran en la tabla 7.2.

TABLA 7.2 Datos de un diseño transversal

Años	1990					
Edad	10	11	12	13	14	15
X en inteligencia	22	23	25	29	30	32
Nº de sujetos	38	39	41	43	42	37

a)

La representación gráfica de la evolución de la inteligencia en las edades estudiadas podría ser la que aparece en la figura 7.1.

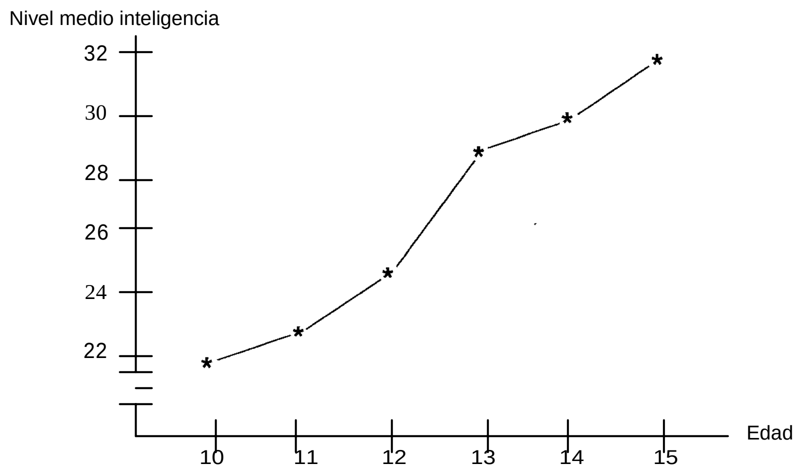


FIG. 7.1 Representación gráfica de la evolución de la inteligencia

c) *Análisis de cohortes*

Los estudios de cohortes combinan algunas características de las estrategias longitudinal y transversal. Como ha podido apreciarse en las tablas 7.1 Y 7.2, los estudios longitudinales tienen por objeto el mismo grupo de sujetos durante el transcurso de los años, y los transversales estudian varios grupos de sujetos en un momento dado. Así, en el estudio de cohortes de la tabla 7.3 se mide la inteligencia de grupos de características similares en cada casilla con respecto a las variables año y edad. Las medidas se toman para cada año y para cada edad implicada, pero no es necesario que los sujetos sean los mismos.

Estos grupos, tal como están representados en la tabla 7.3, se denominan *cohortes* y permiten estudiar tres posibles efectos principales y sus interacciones (Bisquerra, 1989,127):

b)La incidencia de la inteligencia. Para ello se compararían las medias de inteligencia correspondientes a los sujetos incluidos en cada columna.
c)La incidencia de la edad. Comparando las medias de inteligencia que obtienen los sujetos medidos en cada fila.
d)La incidencia de las cohortes. Comparando las medias de inteligencia incluidas en la diagonal.

El estudio de cohortes puede realizarse a partir de datos recogidos en otras investigaciones previas. Para realizar el análisis estadístico y teniendo en cuenta la naturaleza de los datos puede recurrirse a la prueba de ji -cuadrado, al análisis de la varianza o al análisis de la regresión.

TABLA 7.3 Diseño de cohortes. Datos ficticios. En las casillas se han incluido las medias de inteligencia y se ha omitido el número de sujetos para facilitar la comprensión de la tabla

Edades	Años					
	1985	1986	1987	1988	1989	1990
10	21	20	21	23	22	21
11	22	23	22	23	23	22
12	23	24	25	24	25	25
13	26	25	26	27	26	27
14	28	27	28	27	29	29
15	30	29	31	32	31	30

ESTUDIOS DE ENCUESTA

La investigación por encuesta es muy utilizada en el ámbito educativo, probablemente debido a la aparente facilidad y el carácter directo de esta metodología. Esencialmente, la encuesta consiste en formular preguntas directas a una muestra representativa de sujetos a partir de un protocolo o guión previamente elaborado.

Ejemplos de problemas susceptibles de encuesta pueden ser: ¿Qué características tienen los alumnos de secundaria que se presentan a la selectividad universitaria? ¿Cuántos son admitidos? ¿Qué opinan los profesores sobre el tratamiento de la diversidad tal como se plantea en sus centros? La encuesta es muy útil cuando la investigación requiere datos descriptivos que los sujetos pueden proporcionar a partir de su propia experiencia personal.

Concepto y características. La encuesta es un método de investigación basado en una serie de preguntas dirigidas a sujetos que pueden constituir una muestra representativa de una población con el fin de describir y/o relacionar características personales y ciertos ámbitos de información necesarios para responder al problema de investigación, como pueden ser: intereses, motivaciones, creencias, actitudes, intenciones, deseos, percepciones, vivencias y conductas personales o de grupos. Esta información recogida suele referirse al mundo profesional, social o personal, y permite descubrir la frecuencia, la distribución estadística y buscar relaciones entre las variables implicadas, generalizando las conclusiones. Los datos recogidos son necesarios para el investigador y suelen ser analizados en función de otras variables que reflejan características personales como edad, sexo, nivel académico, clase social o profesión.

El método de encuesta suele presentar, por tanto, dos características distintivas.
Por un lado, la recogida de datos se basa en la formulación de preguntas a personas que poseen la información y son capaces de comunicarla, a través de una entrevista personal, por correo o por teléfono. En segundo lugar, el método de encuesta pretende hacer estimaciones de las conclusiones para la población a partir de los resultados obtenidos en una muestra.

Consideraciones sobre el análisis de datos. Con el fin de ilustrar algunas posibilidades de análisis que permite la encuesta estructurada nos basaremos en la hoja de recogida de datos que se reproduce en la figura 7.2.

Señale con una cruz donde corresponda:

*** Sexo:**
A) Femenino __
B) Masculino __

***Edad:**
A) 20-24 __
B) 25-34 __
C) 35-44 __
D) 45-50 __
E) > 50 __

***Estado civil**
A) Soltero __
B) Casado __
C) Viudo __
D) Separado__

*** Las prácticas de mi asignatura se basan en:**
A) Ejercicios libres __ B) Ejercicios dirigidos __

Rodee con un círculo la respuesta más adecuada:

***Creo que un curso de técnicas de estudio para mis alumnos es:**

Poco útil Bastante útil Muy útil

1 2 3 4 5

FIG: 7.2 Extracto del contenido de una encuesta

Supongamos que dicha encuesta se aplica a una representativa de profesores de secundaria. Aunque la muestra debería ser más numerosa, con fines ilustrativos, imaginemos las respuestas de diez sujetos (tabla 7.4). Se han reservado las columnas C1, C2, C3, C4 y C5, respectivamente, para las variables de sexo, edad, estado civil, tipo de práctica y grado de utilidad del curso de técnicas de4 estudio.

TABLA 7.4 Extracto de la recogida de datos correspondiente a diez sujetos (S). Cada columna ha sido encabezada con el símbolo Cn

S	C1	C2	C3	C4	C5
1	A	B	C	B	1
2	A	A	A	A	1
3	B	C	E	A	5
4	A	D	D	B	4
5	B	C	E	A	4

S	C1	C2	C3	C4	C5
6	B	A	B	B	2
7	A	A	A	A	1
8	A	B	B	A	2
9	B	B	B	B	5
10	A	C	E	A	3

A partir de la información recogida, el investigador estará en condiciones de registrar las distribuciones de cada variable. Por ejemplo, la distribución de la variable sexo sería la siguiente:

Sexo	*n*	%
Femenino	6	60
Masculino	4	40
TOTAL	10	100

Con un tamaño muestral adecuado podría constatarse si existen relaciones entre características personales de los sujetos; por ejemplo, entre edad y estado civil:

		20-24	25-34	35-44	45-50	>50	
Estado civil	Solteros	2	1				3
	Casados		2	1			3
	Viudos					3	3
	Separados				1		1
		2	3	1	1	3	10

También se pueden estimar relaciones entre otros aspectos de la información recogida; así ocurriría si quisiera saber si existe relación entre el tipo de prácticas y la importancia concedida a los hábitos de estudio. Dado que la segunda variable se ha medido cuantitativamente, se podrían comparar las medias siguientes en función del tipo de prácticas:

Tipo de prácticas	*Nivel medio grado utilidad*
Ejercicios libres	2
Ejercicios dirigidos	4

Además de tabular las frecuencias y los porcentajes (o proporciones) pueden realizarse análisis como los siguientes (Bisquerra, 1989, 133): análisis cruzado con tablas de contingencia (ji-cuadrado) y análisis multivariable (regresión múltiple y análisis discriminante).

ESTUDIOS OBSERV ACIONALES

Dado que la observación como técnica de recogida de datos será tratada en la parte dedicada al análisis de datos, ahora nos centraremos brevemente en la observación como método de investigación (Anguera, 1985). El método observacional proporciona hechos y puede aportar pautas para configurar teorías (Fox, 1981); sin embargo, para adquirir rango científico, en la acepción empírico-analítica ha de ser intencional o vinculada a teorías e hipótesis (Bunge, 1985) y controlada, es decir, objetiva y comprobable. La investigación observacional es adecuada en contextos educativos que presentan características como las siguientes (Arnau, 1978,24):

- Los sujetos son incapaces de aportar una información verbal. Esta situación puede darse en algunas investigaciones sobre el tratamiento de la diversidad o cuando se estudia la primera infancia.
- Los sujetos no presentan un deseo explícito de informar, debido a reticencias o recelos con respecto al tema investigado.
- En algunas situaciones es probable que los relatos retrospectivos de los sujetos puedan sufrir una distorsión temporal. Por ejemplo, cuando se describe la dinámica seguida en un centro educativo a partir de las percepciones y vivencias de un grupo de exalumnos. En este contexto el método observacinal, basado en datos registrados durante el proceso real de internamiento, puede aportar una información más válida.

Esta modalidad constituye uno de los métodos básicos para el descubrimiento de hipótesis (Undergood y Shaughnessy, 1978), para identificar fenómenos relevantes, para sugerir variables causantes de la acción, registrar conductas que en otros momentos podrían revelarse como efectos, abordar áreas de estudio que no pueden ser tratadas por medio de otras metodologías como la experimental. Por otro lado, su capacidad de operar en situaciones naturales le permite adaptarse mejor al ámbito educativo, resultando imprescindible cuando no es viable ningún otro procedimiento para recoger información. Sin embargo, numerosos factores pueden poner en peligro la validez de sus conclusiones. Así, las dificultades derivadas de la percepción humana, las influencias de la personalidad del observador y de las modificaciones de los sujetos, y la posibilidad de sesgar la focalización de la observación (Anguera, 1978). También debe tenerse en cuenta que algunos aspectos como las intenciones, deseos y motivaciones no son directamente observables.

7.4. MÉTODOS BASADOS EN LA CORRELACIÓN

Muchos fenómenos naturales como la presión atmosférica y la lluvia, la humedad y el crecimiento de las plantas, etc., suelen estar relacionados. De la misma manera, los fenómenos educativos no actúan siempre con independencia, sino que se relacionan y se influyen mutuamente. Para explicar mejor los fenómenos es necesario analizar las relaciones entre las variables implicadas mediante coeficientes de correlación.

A continuación se expondrán algunos conceptos básicos para enmarcar adecuadamente las investigaciones que derivan de la correlación, teniendo en cuenta que el cómputo estadístico se tratará más exhaustivamente en otro volúmen complementario sobre el análisis de datos.

CONCEPTOS BÁSICOS

La correlación es la relación entre dos o más variables, e indica el grado en que tienden a variar conjuntamente, en el mismo o en sentido opuesto. Por ejemplo, ¿cuál es la relación existente entre inteligencia y rendimiento académico? El grado de relación viene expresado por un coeficiente y el tipo de coeficiente que hay que emplear depende de la naturaleza de las variables implicadas. Algunos ejemplos de coeficientes de correlación entre dos variables pueden ser la tau de Kendall, la r de Pearson, tetracórica y biserial. Cuando hay que estimar la relación entre más de dos variables puede convenir la correlación parcial, la múltiple o el coeficiente de concordancia.

Un coeficiente de correlación como larde Pearson aporta dos tipos de información:

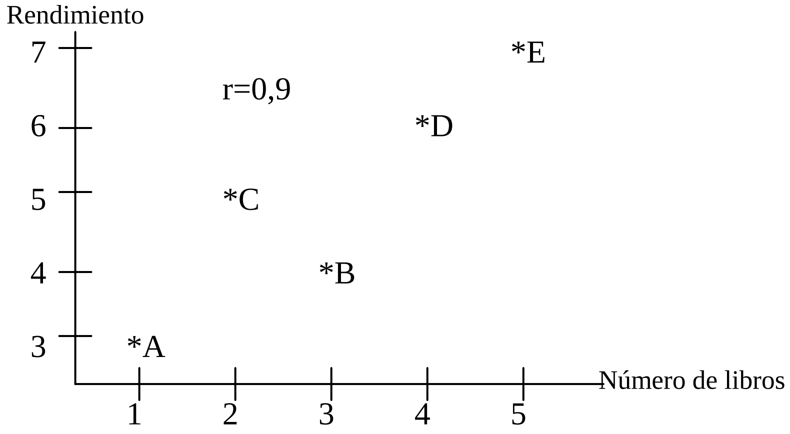
- *Intensidad o grado de relación* entre las variables, expresado a través del coeficiente. Cuando éste posee significación estadística, es decir, cuando es mayor de lo que cabría esperar por efecto del azar, podemos decir que existe dependencia entre las variables implicadas y tiene sentido analizar el tipo de relación.

-*Tipo de relación.* Cuando al aumentar los valores de una variable también tienden a aumentar los valores de la otra, la relación es directa o positiva y la r de Pearson será positiva. En cambio, será negativa o inversa si al aumentar los valores de una variable, los de la otra tienden a disminuir.

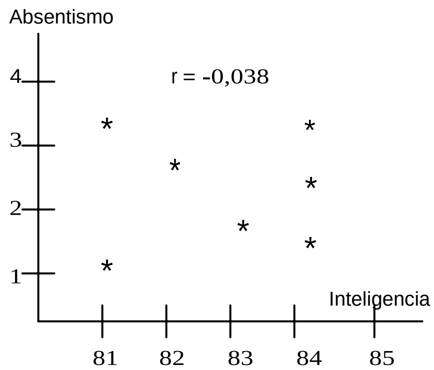
Para relacionar dos variables mediante la correlación se necesitan, como mínimo, dos puntuaciones para cada sujeto, siendo una de cada variable. Los pares de puntuaciones pueden representarse sobre los ejes de coordenadas originando el denominado *diagrama de dispersión*, formado por una nube de puntos. Cada punto corresponde a un par de puntuaciones. La intensidad o grado de relación entre las variables y su dirección quedan visualizados en el diagrama de dispersión y pueden estimarse a partir de un coeficiente de correlación. Supongamos que el número de libros tomados en préstamo de la biblioteca durante cierto período y el rendimiento académico de los alumnos A, B, C, D Y E sean:

Sujetos	*Número de libros* X	*Rendimiento* y
A	1	3
B	2	5
C	3	4
D	4	6
E	5	7

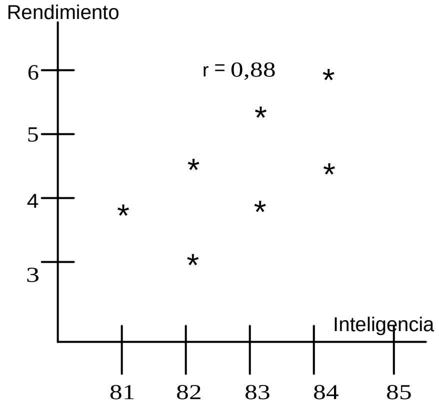
El diagrama de dispersión de las variables anteriores sería el siguiente:



Puede apreciarse que al aumentar el número de libros prestados de biblioteca también tiende a aumentar el rendimiento de los sujetos, y el cálculo estadístico arrojaría un coeficiente r de Pearson de 0,9. Un coeficiente de correlación como la r de Pearson depende de la cantidad de dispersión o varianza alrededor de una línea imaginaria que se ajuste a la nube de puntos. Ejemplos de casos que podrían presentarse son los siguientes:



Ausencia de relación. Independencia

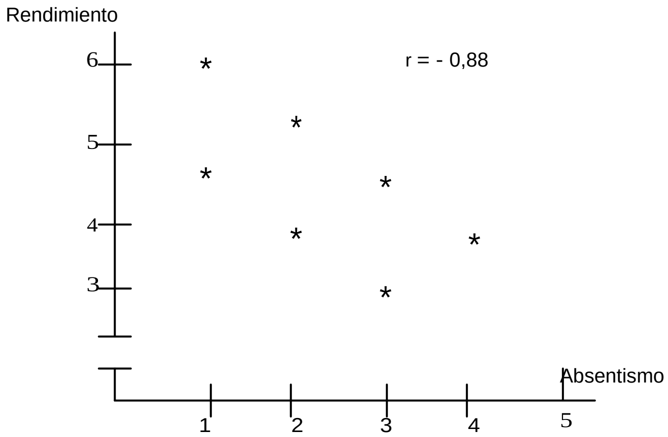


Relación positiva. Dependencia directa

Aunque sería necesario tener en cuenta la significación estadística, la inspección visual de los diagramas permite afirmar que un coeficiente de 0,038 es un grado de

relación casi nulo porque la nube de puntos presenta mucha dispersión, lo que sugiere la existencia de una posible independencia entre las variables. En cambio, un coeficiente de 0,88 sugiere un mayor grado de relación positiva, menos dispersión y una *dependencia directa* entre las variables.

Por otro lado, un coeficiente de -0,88 indica una relación negativa y posiblemente una dependencia inversa que también debería confirmarse estadísticamente. Es decir, a medida que aumenta el absentismo escolar el rendimiento académico tiende a disminuir, como puede apreciarse en el siguiente diagrama de dispersión:



Relación negativa. Dependencia inversa

ESTUDIOS CORRELACIONALES

Para algunos autores el método correlacional no tiene entidad propia, y lo incluyen como una modalidad de la investigación descriptiva (Van Dalen y Meyer, 1983), o bien del método experimental (Mouly, 1978). Sin embargo, muchos autores (Bartolomé, 1978; Borg y Gall, 1983; De la Orden, 1985; García Hoz y Pérez Juste, 1984; Cohen y Manion, 1985, y Keeves, 1988) lo consideran con suficiente entidad propia como para configurar una categoría específica. En esta misma línea, y dado que el método *correlacional* va más allá de la mera descripción, generando con frecuencia *estudios predictivos*, parece oportuno considerarlos de forma diferenciada, ya que el método *predictivo* puede basarse en métodos descriptivos o en la ecuación de regresión. Por otro lado, la complejidad que encierran las técnicas derivadas del método correlacional aconsejan esta opción diferenciada.

Al igual que ocurre con las investigaciones descriptivas, el método correlacional se integra en el denominado modelo *inductivo* de la ciencia (De la Orden, 1985, xv); sin embargo, también puede llegar a la prueba de hipótesis y a buscar explicaciones mediante el estudio de relaciones entre variables.

Para comprender e interpretar las aportaciones del método correlacional vamos a exponer algunos aspectos esenciales como son la varianza compartida entre dos variables, la identificación de la estructura que subyace en una matriz de correlaciones y la correlación como indicador de causalidad entre variables.

Varianza compartida

El coeficiente de *determinación* (CD) indica el porcentaje de variabilidad de una variable que queda explicado o compartido por la otra variable. Este coeficiente se obtiene expresando en porcentaje el cuadrado de la r de Pearson:

$$CD = r^2 \text{ en } \%.$$

Así, en el caso de la inteligencia y el rendimiento académico tendremos:

$$CD = 0{,}88^2 = 0{,}7744 - 77{,}44\ \%.$$

El coeficiente de *alienación* (CA) expresa el porcentaje de variabilidad de una variable que no queda explicado por la variabilidad de otra variable:

$$CA = 1 - r^2 \text{ en } \%.$$

Es decir,

$$CA = 1 - 0{,}88^2 = 1 - 0{,}7744 = 0{,}2256 - 22{,}56\ \%.$$

Gráficamente puede representarse tal como se indica en la figura 7.7.

e)

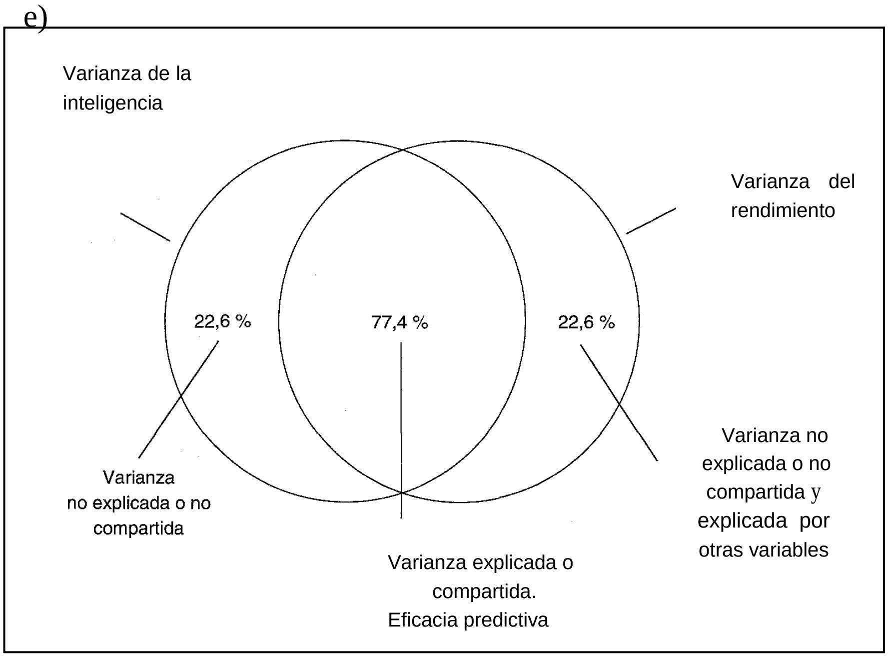


FIG.7.7 Representación gráfica de la correlación entre inteligencia y rendimiento

Un estudio *correlacional* puede orientarse a descubrir o a probar relaciones entre las variables que intervienen en un fenómeno a partir de un coeficiente de correlación. La *matriz de correlaciones* permite la inspección visual de las relaciones existentes entre las variables que intervienen en el fenómeno.

Generalmente, en la investigación correlacional se estudian más de dos variables, y por ello es interesante tener una visión global de la relación mantenida por los distintos pares de variables en una matriz de correlaciones como la siguiente:

Variables	*Inteligencia*	*Absentismo*	*Rendimiento*
Inteligencia	1	-0,038	0,88
Absentismo	–	1	0,88
Rendimiento	-	–	1

A partir de la matriz de correlaciones puede identificarse la estructura subyacente a un grupo de variables a través del denominado *análisis factorial.* Así, en la matriz de correlaciones vemos que la correlación entre inteligencia y rendimiento asciende a 0,88, lo que puede indicar que las dos variables miden o comparten algún aspecto común a ambas, es decir, una misma estructura subyacente, como pudiera ser el razonamiento abstracto.

En la medida que el coeficiente se aparta significativamente de cero y se aproxima a uno, podemos afirmar que las variables están midiendo una característica común o compartida entre ambas. Las variables implicadas ofrecen prácticamente una misma información, puesto que casi están midiendo una misma característica. De forma similar, el número de libros que un grupo de alumnos toman en préstamo de la biblioteca y las horas de lectura que dedican son variables que suelen mantener un alto grado de correlación porque, en definitiva, ambas variables aportan esencialmente una información común que podría denominarse la afición por la lectura. Es decir, la afición lectora puede ser la estructura subyacente de las dos variables anteriores.

En definitiva, el análisis factorial proporciona un medio para simplificar los datos observados o las correlaciones entre variables mediante la reducción del número de variables o dimensiones necesarias para dar la información esencial contenida en una matriz de correlaciones.

Correlación y causalidad

La existencia de correlación significativa entre dos variables es una condición necesaria pero no suficiente para concluir que entre ellas existe una relación de causalidad. Incluso podría ocurrir que una tercera variable, no contemplada en el análisis, fuera el origen de la variación conjunta. Por ejemplo, si en un centro educativo

seleccionamos una muestra representativa de varios cursos escolares, podemos obtener datos como los siguientes: .

Sujetos	*Peso*	*Inteligencia*
A	40	30
B	42	33
C	45	35
..	..	..
..	..	..
..	..	..
y	50	40
Z	53	41

Si se calcula la r de Pearson podría obtenerse un coeficiente significativo, por ejemplo, $r = 0{,}7$. Este coeficiente es aparente y ha sido provocado por la presencia de una tercera variable, que es la edad, relacionada juntamente con el peso y la inteligencia, y que debe introducirse en el análisis: .

Sujetos	(1) *Edad*	(2) *Peso*	(3) *Inteligencia*
A	11	40	30
B	12	42	33
C	13	45	35
	-		. .
	.		
y	14	50	40
Z	15	53	41

Mediante la correlación parcial puede eliminarse o controlarse el efecto de la tercera variable, lo que permite obtener un coeficiente más acorde con la realidad y que en este caso no sería significativo; por ejemplo, 0,12. La representación simbólica del coeficiente sería la siguiente:

$$r_{23}.1 = 0{,}12.$$

Los subíndices indican que el coeficiente expresa el grado de relación existente entre peso e inteligencia controlando estadísticamente la variable edad, es decir, haciéndola constante o eliminando la varianza atribuible a dicha variable.

Aunque la existencia de correlación significativa no sea una condición suficiente para establecer relaciones de causalidad, algunas técnicas novedosas y sofisticadas como el *análisis causal* permiten establecer la plausibilidad de un determinado modelo causal a partir de correlaciones entre variables (Mateo y Rodríguez, 1984, 121-130).

ESTUDIOS PREDICTIVOS

La *predicción* consiste en estimar posibles valores de una variable dependiente o variable criterio a partir de los que toma otra independiente o predictora. El método *predictivo* puede basarse en estudios *descriptivos* y en técnicas como la *ecuación de predicción,* la *correlación canónica* y *la función discriminante* (Bartolomé, 1978,66-67).

Entre los métodos *predictivos* basados en la *descripción* (Bartolomé, 1978,46-60) cabe destacar el método de las *«corrientes»* de Pedro Roselló; el método *proyeccionista,* propuesto por Tusquets; el *análisis de la conducta* utilizado por Escalona y Grace Moore; y las *series temporales* o cronológicas, por su finalidad de analizar el comportamiento de un atributo cuantitativo a través del tiempo. La *situación en miniatura* (Bartolomé, 1978,59) para la predicción del éxito o del fracaso de un sujeto en una determinada profesión o estudios, debería ser decisiva en el ámbito de la orientación escolar y profesional. Se asume que la actuación en la situación en miniatura reflejará la calidad de la actuación en la situación más amplia en la que se desea predecir la conducta.

Técnicas como la *regresión* y la *predicción* son técnicas importantes por su aplicación en el ámbito educativo. La *regresión* consiste en aproximar o hacer «regresar» los puntos de un diagrama de dispersión a una línea recta, con el fin de poder *predecir* valores a partir de la ecuación de dicha recta o ecuación de regresión. La predicción es una conjetura que se formula sobre el valor que tomará una variable a partir de la relación que existe con otra variable.

Para clarificar mejor estos conceptos nos basaremos en un ejemplo. Supongamos que estamos interesados en describir cuáles son los factores determinantes del rendimiento académico (RA), con la intención de llegar a predecirlo. La revisión bibliográfica indica que uno de los múltiples factores que hay que considerar son las horas de estudio dedicadas (HE). Aunque en la investigación deben contemplarse más variables, con fines expositivos, reproducimos un extracto ficticio de la matriz de datos que podría obtenerse con diecinueve sujetos (S) seleccionados al azar de una clase de quinto A:

S	*HE*	*RA*
A	2	1
B	2	3
C	2	5
D	3	1
E	3	3

S	*HE*	*RA*
F	3	5
G	3	7
H	4	1
I	4	3
J	4	5

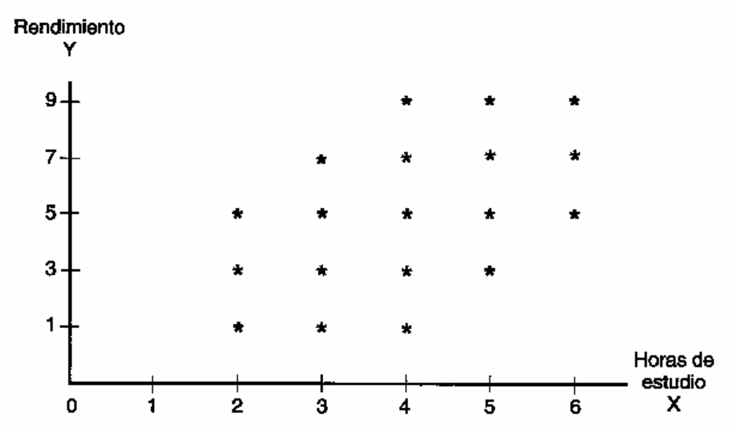
S	*HE*	*RA*
K	4	7
L	4	9
M	5	3
N	5	5
Ñ	5	7

S	*HE*	*RA*
O	5	9
P	6	5
Q	6	7
R	6	9

El diagrama de dispersión muestra la correspondiente nube de puntos formada con los pares ordenados de los 19 sujetos. A partir del diagrama puede verse intuitivamente el significado de *regresión* y *predicción.*

Imaginemos que un sujeto SI asiste a la mencionada clase de quinto A y suele dedicar dos horas de estudio semanales. Gracias al diagrama de dispersión disponemos de cierta información que nos permite intuir cuál será el rendimiento más plausible que obtendrá dicho sujeto. Existen tres sujetos (A, B Y C) que estudian dos horas semanales

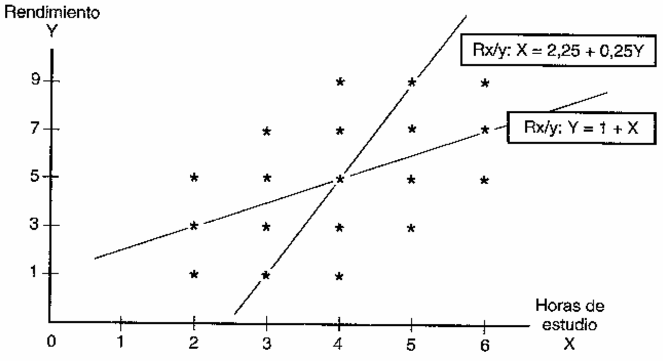
y obtienen, respectivamente, las puntuaciones 1,3 Y 5 en rendimiento. El rendimiento más probable que pronosticaríamos para el sujeto SI sería la media aritmética de las puntuaciones de los sujetos del diagrama que también estudian dos horas a la semana. Así, habría que predecir un rendimiento de 3 puntos. En cierto modo, puede afirmarse que las puntuaciones 1 y 5 han «regresado» a la media del grupo de sujetos que estudian dos horas.



De la misma manera, el rendimiento más probable de un sujeto S2 que dedique tres horas de estudio sería 4, por ser la media de los valores 1, 3, 5 Y 7. Para un sujeto S3 que estudiara cuatro horas cabría predecirle un 5 en rendimiento. Aunque estadísticamente la muestra debería ser más amplia, a partir del diagrama de dispersión podríamos elaborar la siguiente tabla de predicción orientativa:

Horas de estudio (X)	*Rendimiento predecible (Y)*
2	3
3	4
4	5
5	6
6	7

Los pares ordenados que forman la tabla de predicción pueden unirse mediante una recta, que permitiría predecir el rendimiento a partir de las horas de estudio.



La ecuación de esta recta de regresión se denomina *ecuación de predicción* y tiene los siguientes componentes:

$$Y = a + bX$$

y = Variable criterio o variable dependiente. Es la variable que se predice; en nuestro ejemplo es el rendimiento.

X = Variable predictora o variable independiente; en el ejemplo presentado es la variable horas de estudio.

a = Ordenada en el origen. Coincide con el valor de Y cuando X vale cero. En el ejemplo puede apreciarse que a = 1.

b = Coeficiente de regresión, que es la pendiente o tangente de la recta. Coincide con el cambio de Y cuando X cambia en una unidad. También puede obtenerse dividiendo el cateto opuesto por el cateto contiguo. En nuestro caso, b = 1.

En consecuencia, la ecuación de la recta de regresión Ry/x que permite predecir el rendimiento a partir de las horas de estudio será:

$$Y = 1 + X$$

Si quisiéramos predecir las horas de estudio que dedica un sujeto a partir de su rendimiento, seguiríamos un proceso similar, tal como se ha expuesto desde el principio, obteniendo la ecuación:

$$Y = a + bx$$

Sustituyendo,

$$Y = -9 + 4X.$$

Como interesa predecir las horas de estudio, despejaremos la X para obtener la ecuación de la recta de regresión *Rx/y:*

$$X = 2{,}25 + 0{,}25Y$$

Con fines ilustrativos, nos hemos basado en un ejemplo didáctico para obtener la ecuación de la recta de regresión. Sin embargo, la inspección visual no es suficiente en la práctica, por lo que debe acudirse a la estadística informática.

Si se prevé más de una variable independiente, se utiliza la *regresión múltiple,* y cuando hay que predecir más de una variable dependiente se recurre a la correlación canónica. Por ejemplo, en una investigación podemos estar interesados en predecir la comprensión y la velocidad lectoras como dimensiones de la eficiencia lectora, a partir de variables predictoras como la inteligencia, horas de estudio y motivación.

Otras técnicas como el análisis discriminante ofrecen la predicción de la pertenencia a un grupo. Dados grupos de sujetos, por ejemplo, adaptados-inadaptados y delincuentes-no delincuentes, se predice a qué grupo pertenecerá un sujeto en función de otras variables predictoras.

7.5. POSIBILIDADES Y LÍMITES

La metodología ex-post-facto suele ser apropiada en las primeras aproximaciones a una área problemática por su carácter exploratorio, pues facilita la generación de hipótesis, mientras que el método estrictamente experimental es más utilizado en las etapas finales, para la contrastación de los supuestos previos (Mayntz, 1983, y Mouly, 1978).

Cuando se quieren estudiar relaciones de *causalidad,* la mayor validez interna puede conseguirse al utilizar las metodologías *experimental y cuasiexperimental,* siempre que las variables sean susceptibles de manipulación. Sin embargo, a causa de la complejidad y la naturaleza de los fenómenos sociales, el educador no siempre puede seleccionar, controlar y manipular todos los factores necesarios para determinar la relación de causalidad a través de un experimento. De ahí que, para explorar o confirmar relaciones de causalidad, en muchas situaciones educativas sólo sea viable una primera aproximación con un planteamiento ex-post-facto. Así, puede recurrir se al método *comparativo causal* o *selectivo comparativo* y al *correlacional,* del que derivan los modelos causales. Ahora bien, como veremos posteriormente, en la investigación ex - post - facto las relaciones significativas entre variables son condición necesaria pero no suficiente para poner de manifiesto las relaciones de *causalidad.* También hay que tener en cuenta que los estudios *correlacionales* y los *modelos causales* son más adecuados cuando las variables implicadas son cuantitativas y pueden medirse como mínimo con una escala de intervalo.

Como conclusión final podemos decir que la metodología ex-post-facto es apropiada y conveniente desde la doble perspectiva científica y educativa. Ofrece conocimientos necesarios para el desarrollo de la educación como disciplina científica y es imprescindible en muchos ámbitos educativos. Sin embargo, por sí sola no es suficiente para desarrollar el conocimiento científico. La ciencia precisa de métodos complementarios de contrastación que ofrezcan más garantías de control, con el fin de

validar las hipótesis generadas a partir de los métodos descriptivos. En consecuencia, desde la perspectiva estrictamente científica las aportaciones del método ex-post - facto deben complementarse con el método experimental.

Los métodos *descriptivos* y *correlacionales* están más asociados al método teórico de la inducción y permiten llegar a generalizaciones empíricas a través del establecimiento de regularidades y relaciones entre los datos observados, mientras que el método experimental opera con el método deductivo, tratando de contrastar hipótesis derivadas deductivamente de una teoría (De la Orden, 1985, xv).

Algunas técnicas de análisis basadas en la *correlación,* como la *ecuación de regresión* o el *análisis factorial* (Kerlinger, 1985,461-472), o que siguen el *modelo predictivo,* como la *correlación canónica* (Peris, 1986,5-19) Y la *función discriminante* (Bartolomé, 1978,37-49, Y Villanueva, 1984,5-18) tienen muchas posibilidades en educación. En efecto, dada la gran multiplicidad de variables que intervienen en la realidad educativa son muy valiosas las aportaciones del método *correlacional,* pues ofrece la posibilidad de medir y relacionar gran número de variables, puede ser utilizado en contextos naturales y deriva con facilidad en *predicciones.* Así, adquieren especial relevancia e interés en el campo de la orientación escolar y en el diagnóstico pedagógico.

El mayor inconveniente de estos métodos, que provenía de no poder establecer relaciones de causalidad, está siendo superado a través de la aplicación de técnicas estadísticas sofisticadas tales como el *análisis causal* (Mateo y Rodríguez, 1984, 103-130, y Fuentes, 1986,35-48).

La incidencia de estos métodos en el ámbito educativo es enorme, especialmente en áreas como formación del profesorado, orientación escolar, sociología de la educación, didáctica, tecnología educativa y organización escolar. Describen situaciones, actitudes e interacciones en el ámbito educativo, y puesto que sus diseños también pueden utilizarse. en el contexto de la investigación evaluativa, aportan información para orientar la política educativa y la toma de decisiones.

PARTE III

METODOLOGÍAS DE INVESTIGACIÓN EDUCATIVA DE PERSPECTIVA HUMANÍSTICA-INTERPRETATIV A

La perspectiva humanístico interpretativa (cualitativa, etnográfica, naturalista) se constituye como una metodología alternativa a la empírico-analítica de corte positivista (Marshsall y Rossman, 1989) que hemos tratado en los capítulos anteriores.

Se orienta a describir e interpretar los fenómenos educativos y se interesa por el estudio de los significados e intenciones de las acciones humanas desde la perspectiva de los propios agentes sociales. Desde esta perspectiva se aborda el mundo personal de los sujetos (cómo interpretan las situaciones, qué significado tienen para ellos) no observable directamente ni susceptible de experimentación.

La metodología que ~ esta perspectiva, por lo general, sigue un enfoque holístico-im!uctivo-idjográfico, es decir, estudia la realidad 0en su globalidad, sin fragmentarla y contextualizando, sigue la vía inductiva: las categorías, explicaciones e interpretaciones se elaboran partiendo de los datos y no de teorías previas, y se centra en las peculiaridades de los sujetos más que en el logro de leyes generales.

Los diseños son de naturaleza flexible y adoptan un enfoque progresivo Los métodos están al servicio del investigador y no a la inversa. Como técnicas de recogida de datos tiende a utilizar estrategias de tipo cualitativo, como la entrevista, la observación participante, notas de campo, análisis de documentos, etc.

Desde esta perspectiva se cuestionan los criterios científicos de validez, fiabilidad y objetividad propios de la metodología empírico-analítica, y se adoptan otros como procesos de triangulación y contrastación intersubjetiva.

En el capítulo 8 se exponen brevemente algunas escuelas de pensamiento, como la fenomenología, el interaccionismo simbólico o la etnometodología, que han aportado los principios que fundamentan la metodología interpretativa. La investigación etnográfica y el estudio de casos se consideran como modalidades de dicha metodología.

CAPÍTULO 8

Perspectivas de orientación interpretativa

En esta perspectiva se engloba una familia de métodos unidos por su orientación a describir. e interpretar la realidad social y su oposición al enfoque positivista orientado a explicar y descubrir las regularidades de los fenómenos. De entre estos métodos cabe señalar la observación participante y etnografía (Adler y Adler, 1987), el trabajo de campo (Manning, 1987), la investigación etnográfica posmodema (Tumer y Bruner, 1986), los estudios naturalísticos (Lincoln y Guba, 1985), la entrevista creativa (Douglas, 1985), el estudio de caso (Stade, 1986), los trabajos interpretativos, hermenéuticos y fenomenológicos (Gadamer, 1975), los estudios de enfoque cultural (Hall, 1980) Y el interaccionismo interpretativo (Denzin, 1989).

De manera muy sucinta exponemos algunas escuelas que por su orientación interpretativa han tenido y siguen ejerciendo gran influencia en la perspectiva de investigación que venimos denominando humanístico-interpretativa. Haremos referencia solamente a la fenomenología, la etnometodología y el interaccionismo simbólico.

8.1 FENOMENOLOGÍA

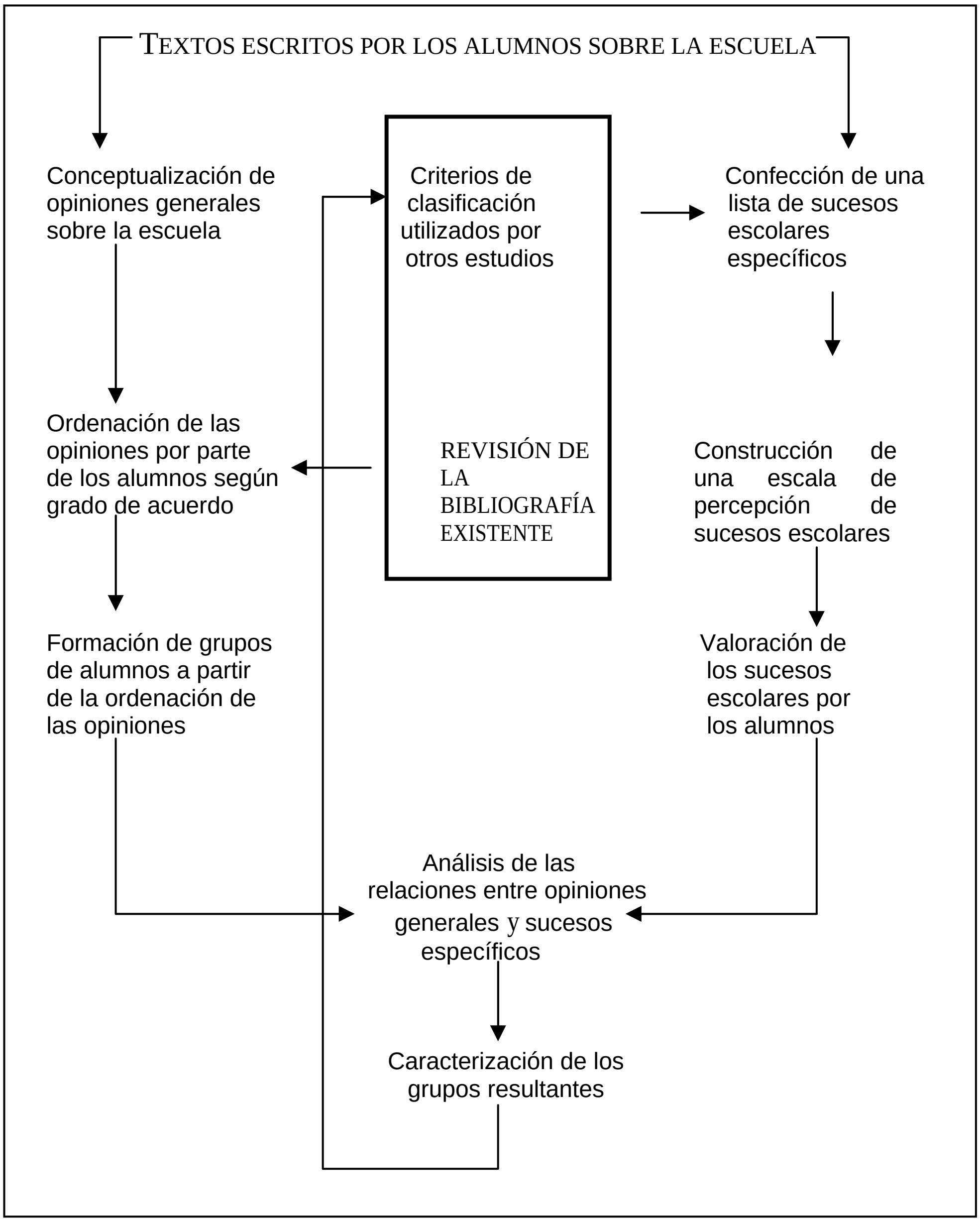
En un sentido general, por fenomenología se entiende la escuela de pensamiento o filosofía creada por E. Husserl (1859-1938) en la primera parte del siglo xx, cuyos principios son esenciales para la comprensión de la corriente interpretativa.

Como notas más destacadas que aporta a la metodología interpretativa cabe señalar: *a)* la primacía que otorga a la experiencia subjetiva inmediata como base del conocimiento; *b)* el estudio de los fenómenos desde la perspectiva de los sujetos, teniendo en cuenta su marco referencial; e) su interés por conocer cómo las personas experimentan e interpretan el mundo social que construyen en interacción.

En este sentido la fenomenología introduce un nuevo enfoque como vía de conocimiento de los fenómenos. Si bien son escasos los estudios de este tipo en el ámbito educativo, los realizados se orientan a la búsqueda de los significados que los sujetos otorgan a sus experiencias como miembros de las instituciones.

El cuadro 8.1 recoge el diseño aplicado por los profesores San Fabián y Corral en un modelo de evaluación fenomenológica de la escuela.

CUADRO 8. 1 Diseño de la investigación



Para realizar dicho estudio los autores se sirvieron de dos instrumentos: una lista de opiniones sobre la escuela y una escala de percepción de sucesos escolares, elaborados a partir de los escritos de los alumnos de séptimo y octavo en una hoja de respuestas cuyo encabezamiento decía: «Escribe una carta a un/a amigo/a tuyo/a contándole lo que opines sobre el colegio».

8.2 ETNOMETODOLOGÍA

La etnometodología es una corriente sociológica de pensamiento que nació en los años sesenta en los Estados Unidos en tomo a la obra de Harold Garfinkel (1967) *Studies in Ethnomethodology,* cuyos orígenes deben situarse en tomo ala tradición fenomenológica.

La importancia teórica y epistemológica de la etnometodología viene dada por el hecho de que adopta planteamientos que rompen radicalmente con la postura de la tradición sociológica. «Más que una teoría constituida, es una perspectiva de búsqueda, una nueva postura intelectual» (Coulon, 1988,9).

La etnometodología se centra en el estudio de los métodos empleados por las personas para dar sentido a sus prácticas sociales cotidianas. El etnometodólogo trata de entender cómo las personas «emprenden la tarea de ver, describir y explicar el orden en el mundo en que viven» (Zimmerman y Wideder, cito por Taylor y Bogdan, 1970, 27). Pone el énfasis en el análisis cualitativo detallado de las pautas de interacción social, en la manera en que la gente crea y construye sus formas de vida, el orden y las reglas sociales.

Parte de los supuestos de que el orden social es construido por los actores sociales y que los eventos sociales son resultado de las interacciones de los sujetos implicados en los mismos. El orden social no depende de las normas o presiones sociales, sino que deviene como resultado de la implicación de las personas en lograrlo (McNeill, 1985).

La etnometodología ha desarrollado varios métodos, que Garfinkel denomina *procedimiento de fractura,* para buscar cómo las reglas y normas sociales son logradas por los actores sociales: el experimento disruptivo, que consiste en interrumpir deliberadamente una rutina social establecida y ver qué ocurre; y el análisis conversacional, que analiza los significados de las conversaciones teniendo en cuenta el contexto (McNeill, 1985).

Esta modalidad de investigación ha sido aplicada al ámbito educativo por H. Mehan (1978) para estudiar las reglas y el orden social de las aulas.

8.3 INTERACCIONISMO SIMBÓLICO

Esta escuela surge como enfoque alternativo a los estudios sociológicos de los años cuarenta y cincuenta de corte conductista. Fue desarrollada por H. Blumer (1969), aunque sus orígenes pueden situarse en los trabajos de G. H. Mead (1938), Ch. H. Cooley (1902), W.1. Thomas (1931) y los estudios de la escuela de Chicago (Taylor y Bogdan, 1986).

Sus principios comparten la perspectiva fenomenológica y se basan en el supuesto de que la experiencia humana está mediada por la interpretación (Blumer, 1982). Los seres humanos actúan sobre las cosas en función del significado que tienen para ellos. Los interaccionistas simbólicos se interesan por describir los procesos de interpretación como instrumento de comprensión de los significados de las acciones humanas.

El interaccionismo simbólico atribuye importancia primordial a los significados sociales que las personas atribuyen a los objetos que les rodean mediante símbolos (Blumer, 1982). Según el mismo autor, el interaccionismo simbólico se basa en tres premisas básicas: *a)* la actuación de las personas sobre las cosas, e incluso sobre las personas, depende de lo que significan para ellas; *b)* los significados son el resultado de la interacción; y c) la interpretación genera y modifica los significados.

Este enfoque ha sido aplicado al ámbito educativo por H. Becker (1961), en *Boys in White,* donde aborda el tema de cómo los alumnos, durante su carrera escolar, resuelven sus problemas inmediatos con los profesores y las tareas asignadas; y por Ph. Cusick (1973) en su obra *Inside High School,* donde hace un estudio de la interacción de los alumnos en un contexto escolar.

CAPÍTULO 9

Investigación etnográfica

9.1 PERSPECTIVA HISTÓRICA

Los orígenes de la moderna etnografía se sitúan a finales del siglo XIX y comienzos del XX, en los trabajos de los antropólogos sobre el modo de vida de tribus bajo dominio del Imperio británico, conscientes de la importancia que tenían las culturas de los «pueblos primitivos» y que podían perderse por la influencia de la cultura europea (MacNeill, 1985,56).

El primer tercio del siglo xx vivió el desarrollo de la etnografía llevado a cabo por la antropología cultural. Los antropólogos se interesaron por la cultura de los pueblos primitivos y marcharon a vivir entre ellos para estudiar sus sociedades «desde dentro». La obra de Malinowski (1922) sobre los habitantes de las islas Trobriand marca un nuevo enfoque al abordar el estudio de estas culturas, que hasta ahora se había realizado «desde fuera». Para este autor el único modo efectivo de comprender la forma de vida de estas personas era «ir y vivir» entre ellos por un período de tiempo, aprendiendo su lengua y llegar a ser aceptado como miembro. El trabajo lo realiza de manera descriptiva, sistemática y técnica, combinando la observación participante con la entrevista.

En los años cincuenta se desarrolló en el seno de la antropología el estudio de la cultura dentro del ámbito educativo, lo que dio lugar a la antropología educativa (institucionalizada por los profesores estadounidenses S. Kimball y G. Spindler) y que estudia la influencia de la cultura en la educación: los procesos de enculturización (las formas en que los niños y adolescentes se introducen en las prácticas y creencias de los grupos sociales) y aculturización (los efectos que produce la falta de escolaridad) (Goetz y LeCompte, 1988).

A partir de los años setenta la antropología educativa diversifica sus temas de estudio y métodos; no sólo estudia los aspectos culturales de la educación (Erickson, 1977), sino que se extiende a otros aspectos como el lenguaje (Cazden y otros, 1972; Cook-Gumperz, 1976), los sistemas simbólicos mediante los que la cultura se transmite, mantiene y transforma; los sistemas conceptuales tácitos (como la interrelación y participación a través de las que se crean y expresan los procesos culturales en los escenarios educativos (Philips, 1972), los sistemas de evaluación educativa (Fetterman, 1979).

Juntamente con la antropología, aunque algo más tarde, la sociología contribuye a desarrollar la etnografía en las ciencias sociales. En las décadas de los veinte y treinta la escuela de Chicago aplica las técnicas antropológicas al estudio de estilos de vida de grupos marginados (bandas juveniles, gánsteres, emigrantes) en contextos urbanos y emplea constructos sociológicos (sociedad, socialización). En el campo de la educación son pioneros los trabajos de Becker y otros (1961) y Smith y Geoffrey (1968), quienes usan los métodos antropológicos para el estudio de los procesos del aula. En la década de los setenta, estos estudios irrumpen con fuerza en la investigación educativa (Cazden, 1972; Spradley, 1972; Erickson, 1975; Wilson, 1977. En los sesenta, el grupo de Edimburgo (Arkinson, Harnilton, Stubbs, Reid) lleva a cabo estudios etnográficos, siendo pionero en el desarrollo de la etnografía educativa en el Reino Unido.

9.2 CONCEPTO Y CARACTERÍSTICAS

La etnografía es considerada como una modalidad de investigación de las ciencias sociales que surge de la antropología cultural y de la sociología cualitativa, y se inscribe en la familia de la metodología interpretativa/cualitativa. Se erige como modelo alternativo a la investigación tradicional utilizada por los científicos sociales para estudiar la realidad social.

En sentido amplio, la expresión *investigación etnográfica* viene utilizándose como sinónimo de *investigación cualitativa, trabajo de campo, observación antropológica,* etc. Se define como una descripción o reconstrucción analítica de los escenarios y grupos culturales intactos (Spradley y McCurdy, 1972); como un planteamiento de hacer investigación naturalista, observaciona1, descriptiva, contextual, no limitado de antemano y en profundidad (Harnmersley y Atkinson, 1983); como el arte y ciencia de describir un grupo o cultura (Fetterman, 1989).

En su sentido literal significa descripción de un modo de vida, de una raza o grupo de individuos (Woods, 1987, 12); una monografía detallada de los modos de vida de los *etnoi,* término griego que designa a los «otros», a los que no eran griegos (Erickson, 1986).

La investigación etnográfica esencialmente consiste en una descripción de los eventos que tienen lugar en la vida del grupo, con especial consideración de las estructuras sociales y la conducta de los sujetos como miembros del grupo, así como de sus interpretaciones y significados de la cultura a la que pertenecen (Taft, 1988; 59).

Cuando este estilo de investigación se utiliza para estudiar los fenómenos o realidad educativa se denomina *etnografía educativa.* Se entiende como una descripción detallada de las áreas de la vida social de la escuela (Woods, 1987; 24). Ofrece un estilo de investigación alternativo para describir, explicar e interpretar los fenómenos educativos que tienen lugar en el contexto de la escuela.

El uso de la investigación y métodos etnográficos ha ido incrementándose entre los investigadores educativos durante las dos últimas décadas, adquiriendo una gran importancia para el estudio de contextos educativos, siendo para Lutz (1981) una línea de investigación prometedora en el ámbito de la educación.

El enfoque etnográfico intenta describir la totalidad de un fenómeno (grupo social, aula, etc.) en profundidad y en su ámbito natural, y comprenderlo desde el punto de vista de los que están implicados en él. Como rasgos más destacados que configuran

la naturaleza de la etnografía, aunque no exclusivos de la misma, cabe señalar los siguientes:

a) Su carácter holista: describe los fenómenos de manera global en sus contextos naturales, aceptando el escenario complejo que encuentra y la totalidad, como elementos básicos.

b) Su condición naturalista: el etnógrafo estudia las personas en su hábitat natural. Observa, escucha, habla, anota las historias de vida y evita las formas controladas. e) Usa la vía inductiva: se apoya en las evidencias para sus concepciones y teorías, y en la empatía y habilidad general del investigador para estudiar otras culturas.

d) Su carácter fenomenológico o émico: los significados se estudian desde el punto de vista de los agentes sociales.

e) Los datos aparecen contextualizados: las observaciones se sitúan dentro de una perspectiva más amplia.

f) Libre de juicios de valor: el etnógrafo evita emitir juicios de valor sobre las observaciones.

Knapp (cit. por Cook y Reichardt, 1986, 172) señala como elementos que caracterizan la investigación etnográfica: *a)* un acceso al escenario inicialmente exploratorio y abierto a las contingencias del problema de investigación; *b)* una intensa implicación del investigador en el entorno social que estudia como observador y participante; *c)* el empleo de variedad de métodos y técnicas cualitativas, siendo las más frecuentes la observación participante, la entrevista informal y el análisis documental, llevar un registro cuidadoso de lo que ocurre y anotar toda clase de evidencia documental; *d)* una tentativa explícita de comprender los acontecimientos en términos de significado; *e)* un marco interpretativo que subraye el importante papel del contexto; y *f)* una elaboración de los resultados de la investigación en forma descriptiva.

En sentido amplio, la etnografía tiene como fin el estudio sociocultural o estilo de vida de la sociedad, describiendo las creencias y prácticas del grupo, mostrando cómo las diversas partes de la comunidad contribuyen a crear la cultura como un todo unificado y consistente (Jacob, 1987). «Se interesa por lo que la gente hace, cómo se comporta, cómo interactúa. Se propone describir sus creencias, valores, perspectivas, motivaciones y el modo en que todo ello se desarrolla o cambia dentro del grupo y desde las perspectivas de sus miembros. Lo que cuentan son su significados e interpretaciones» (Woods, 1987, 18).

Por lo que concierne a la etnografía educativa, su objetivo es aportar valiosos datos descriptivos de los ámbitos o contextos educativos, actividades y creencias de los participantes en los escenarios educativos (Goetzy LeCompte, 1988,41), describir las diversas y complejas perspectivas y actividades de los profesores y alumnos con vistas a obtener explicaciones para descubrir patrones de comportamiento (Harnmersley, 1985), proveer un análisis de la escuela como un sistema dinámico de relaciones sociales.

En el ámbito de la formación del profesorado se ha usado como medio de unión de la teoría con la práctica, como ayuda para observar y analizar la práctica educativa (Gitlin y Teitelbaum, 1983); como medio de formación reflexiva y autocríticamente o como instrumento para mejorar la competencia profesional (Isaac, cit. en Pollard, 1983).

9.3 EL PROCESO ETNOGRÁFICO

Las fases de la investigación etnográfica están menos definidas que las de otros métodos de investigación, pero se pueden identificar algunas fases.

Denzin (1978) señala que en la investigación etnográfica deben considerarse estas siete áreas de decisión: 1) el foco y finalidad del estudio y las cuestiones que aborda; 2) el modelo o diseño de investigación utilizado y las razones de su elección; 3) los participantes o sujetos del estudio, el escenario y contexto(s) investigados; 4) la experiencia del investigador y sus roles en el estudio; 5) las estrategias de recogida de datos; 6) las técnicas empleadas en el análisis de datos: y 7) los descubrimientos del estudio: interpretaciones y aplicaciones.

Para Goetz y LeCompte (1988, 172) son cuatro las fases de la investigación etnográfica. En la primera fase se plantean las cuestiones relativas a la investigación y marcos teóricos preliminares y se selecciona un grupo para el estudio. En la segunda se aborda el acceso al escenario, la elección de los informantes clave, el comienzo de las entrevistas y las técnicas de recogida de datos y de registro. En la tercera se aborda la recogida de datos; y, por último, en la cuarta se lleva a cabo el análisis e il1terpretación de los datos.

Ilustramos el proceso etnográfico reseñando el modelo llevado a cabo en las universidades norteamericanas de Utah y Wisconsin-Madison en la formación de profesores. El procedimiento sigue los siguientes pasos:

a)Un seminario inicial para discutir la naturaleza de la investigación etnográfica.

b) Readings sobre investigación etnográfica (por ejemplo, Patton. 1975; Willis, 1978).
c) Discusiones-debate sobre estudios etnográficos (por ejemplo, Jackson, 1968; Spradley y McCurdy, 1972; Spindler, 1974).
d) Un estudio etnográfico sobre un aula que incluye:

- la historia interna, reglas y filosofía establecida, arquitectura y facilidades de la escuela;
- localización de materiales, reglas establecidas y no establecidas, atmósfera general de la clase;
- la conducta de un estudiante y las interacciones verbales y no verbales con los otros miembros (docentes, estudiantes) durante un día entero;
- las actitudes de los docentes hacia su trabajo;
- otros aspectos de la escuela que los estudiantes eligen como foco de observación;
- un informe: con resultados, áreas necesitadas de ser más exploradas e implicaciones para la práctica educativa;
- discusión de los resultados por todo el grupo de alumnos (Gitlin y Teitelbaum, 1983).

9.4 TÉCNICAS Y ESTRATEGIAS METODOLÓGICAS

La etnografía como modalidad de investigación utiliza múltiples métodos y estrategias. El diseño etnográfico supone una amp1ia combinación de técnicas y recursos metodológicos; sin embargo, suelen poner mayor énfasis en las estrategias

jnteractivas: la observación participante, las entrevistas formales e informales, los instrumentos diseñados por el investigador y el análisis de toda clase de documentos.

Es evidente que también puede utilizarse una variada tecnología compuesta de magnetófonos, dictáfonos, procesadores de palabras, ordenadores *(laptop* y *desktop)*, bases de datos *(hypertex)*, cámaras fotográficas y equipos de vídeo; todas las herramientas del etnógrafo son extensiones del instrumento humano, ayudas a la memoria y visión.

1) *La observación participante.* El investigador etnográfico combina la observación con la participación. El etnógrafo es el agente principal de la investigación; observa las pautas de conducta y participa en la cultura que está siendo observada, El grado de participación y observación varía de un estudio a otro; algunas veces el etnógrafo asume el rol de observador y en otras de participante; en ambos roles permanece atento a lo que ocurre, reflexionando y recogiendo registros de campo.

La observación participante ha sido siempre el método central de los etnógrafos. A menudo lo complementan con otras fuentes, en especial con entrevistas informales (McNeill, 1985,58).

2) La entrevista informal. Es otra técnica utilizada por los etnógrafos en el trabajo de campo. Se obtiene información participante. Su objetivo es mantener a los participantes hablando de cosas de su interés y cubrir aspectos de importancia para la investigación en la manera que permita a los participantes usar sus propios conceptos y términos (Whyte, 1982, 112).

La entrevista informal puede variar desde discusiones casuales mientras participan en una actividad, a entrevistas abiertas o discusiones en profundidad que seleccionan sujetos llamados *informantes clave.*

3) Materiales escritos (documentos). La utilización de materiales escritos o impresos constituye un apoyo útil a la observación (Woods, 1987). Los materiales de este tipo que más se usan son los documentos oficiales, los personales y los cuestionarios. Los materiales escritos han llegado a ser considerados como instrumentos cuasiobservacionales (Woods, 1987). En cierto modo reemplazan al observador y al entrevistador en situaciones inaccesibles.

- *Documentos oficiales.* Incluyen registros, horarios, actas de reuniones, programaciones, planificaciones y notas de lecciones, registros personales de alumnos, manuales escolares, periódicos y revistas, grabaciones escolares, archivos y estadísticas, tableros de anuncios, exposiciones, cartas oficiales, libros de ejercicios, exámenes, fichas de trabajo, fotografías (Woods, 1987).
- *Documentos personales.* Como diarios, ejercicios de escritura creativa, el cuaderno borrador de los alumnos, graffiti, cartas y notas personales.
- *Cuestionarios.* Sin ser populares entre los investigadores etnográficos, se les atribuye utilidad como medio de recogida de información de muestras más amplias que las que se pueden obtener por medio de la entrevista; como punto de partida para el uso de métodos más cualitativos, y como subsidiarios de técnicas interpretativas (por ejemplo, las sociométricas) (Woods, 1987).

9.5 DIFICULTADES QUE PLANTEA LA ETNOGRAFÍA

Vemos la necesidad de hacer algunas consideraciones sobre el uso de la etnografía en lo que concierne a aspectos relacionados con el acceso al escenario, el rol del

investigador, los informadores clave, la subjetividad, que precisan de una atención singular.

a) *Acceso al escenario.* Para Erickson (1986), la entrada o acceso al escenario es un fenómeno complejo que en consecuencia exige tacto y actitud diplomática. Tanto si el observador es aceptado como si despierta recelo, es el período que resulta más incómodo y desagradable. Por ello se hacen recomendaciones y reflexiones sobre cómo conducirse.

Los elementos que tienen más importancia en el acceso a los grupos y entrada en el escenario de la investigación son la flexibilidad del investigador y su sensibilidad a respuestas de los participantes. Requiere que el investigador presente los fines de su trabajo de la forma más sencilla y tentativa posible (Goetz y LeCompte, 1988, 108).

La entrada del etnógrafo al escenario va acompañada de una primera etapa de contacto, llamada *de vagabundeo* (LeCompte, 1969), que proporciona los datos de base para ir explorando la situación y superando las dificultades que van surgiendo. El *vagabundeo* consiste en reconocer el terreno, familiarizarse con los participantes y documentarse sobre la situación. En las primeras etapas de contacto con el escenario es importante despertar confianza y mostrarse como una persona de valor e integridad. La mejor manera de conseguirlo es tener un proyecto honesto, que sea útil para la mejora de la educación (Goetz y LeCompte, 1988) .

. . El contacto con las fuentes de datos se puede entablar de manera formal o informal. El contacto formal supone un acercamiento a través de los canales oficiales y, a menudo, profesionales. El contacto informal se basa en la utilización de redes de relaciones personales. Ambos contactos pueden iniciarse tanto personalmente como por escrito. Por lo general, el contacto informal y directo es más eficaz. Todo ello, no obstante, dependerá de la situación (Goetz y LeCompte, 1988, 107).

Los modos más comunes de acceso a los escenarios suelen ser: *a)* a través de relaciones recíprocas: aportaciones, ayudas y beneficios que se intercambian; *b)* introducirse en el escenario si los participantes no están abiertos a la observación y no quieren colaborar; y c) acudir a un avalador reconocido o autoridad.

b) *Informantes o actores clave.* La etnografía utiliza más el término de *participantes* que el de *sujetos*, con lo que reconoce el papel activo que éstos desempeñan en la investigación. Es importante que el etnógrafo mantenga buenas relaciones con todos los informantes para el éxito de la investigación, pero éstas se hacen necesarias con los informadores clave (término usado por la antropología). Una forma de conseguirlo es implicarlos en el proceso de la investigación (Erickson, 1986).

Los actores clave son participantes que ayudan a tener una mayor comprensión del escenario y de la situación. «Constituyen fuentes primarias de información, actuando de «protectores» del investigador: lo presentan a otras personas, responden por él, le dicen cómo debe actuar y le informan de cómo es visto por los demás participantes» (Taylor y Bogdan, 1986).

Los informantes clave son participantes que poseen conocimientos, *status* o destrezas comunicativas especiales y que están dispuestos a cooperar con el investigador (Zelditch, 1962). Ayudan al investigador a penetrar las barreras que se interponen en su camino. Desempeñan un cometido importante para acceder a determinados subgrupos y personas que por otra vía serían de difícil alcance.

Recurrir a actores clave conlleva cierto riesgo; sus perspectivas y opiniones pueden estar distorsionadas o sesgadas; se deben elegir con cuidado si se pretende que tengan un nivel adecuado de representatividad respecto del grupo completo de informantes clave. Patton (1987) recomienda que la información obtenida de los informantes clave sea claramente especificada y diferenciada como tal en las notas de campo.

c) *El rol del etnógrafo.* El etnógrafo es, primero y ante todo, el «instrumento esencial de la investigación». Los etnógrafos son recolectores de datos, «observadores, narradores y escritores» (Wolcott, 1975, 115). Admite su subjetividad y la de los participantes. Su trabajo requiere relacionarse e interactuar en los escenarios educativos de forma natural. El rigor de la investigación radica más en la actuación del propio investigador que en los instrumentos que utiliza.

El rol del etnógrafo llegó a considerarse como valor científico cuando la estancia en una cultura extraña planteaba múltiples problemas. En el devenir del estudio adopta distintos roles (observador, participante, investigador) y *status* dentro de la comunidad. La implicación con los participantes le plantea diversos problemas, siendo el más evidente el de cómo ser objetivo sin perder su perspectiva como investigador.

El etnógrafo no sólo requiere adaptarse a los distintos roles que ha adquirido o le han sido asignados en el escenario, sino que también debe tener la capacidad de cambiar eficazmente de rol cuando las circunstancias lo precisen. A la habilidad de comunicarse con otros grupos culturales y de adaptarse a su hábitat natural se *denomina franqueo de fronteras;* su dominio se considera decisivo para el éxito de un proyecto de investigación de esta naturaleza.

9.6 ALGUNAS ORIENTACIONES SOBRE LOS ESTUDIOS ETNOGRÁFICOS

A pesar de las dificultades que conlleva precisar normas y procedimientos para llevar acabo los estudios etnográficos, Patton (1987) señala las siguientes orientaciones:

a) Ser descriptivo al tomar las notas de campo.
b) Hacer acopio de gran variedad de información procedente de distintas perspectivas.
e) Triangular y efectuar validaciones cruzadas recogiendo diferentes tipos de datos, a través de observaciones, entrevistas, documentos, archivos y fotografías.
f) Utilizar citas y relatos literales en el lenguaje empleado por los participantes.
g) Seleccionar con cuidado los informantes clave y tener en cuenta que sus perspectivas son limitadas.
j) Ser consciente de las diferentes etapas del trabajo de campo:

- Ganarse la confianza y buscar el *rapport* en la fase de acceso. Recordar que el observador también está siendo observado.
- Estar alerta y disciplinado durante la fase rutinaria intermedia durante la recogida de datos.
- Centrarse en la elaboración de una síntesis útil a medida que se acerca el fin del trabajo de campo.
- Ser disciplinado y concienzudo durante la toma de las notas de campo durante todas las fases de la investigación.

g) Implicarse tanto como sea posible en la realidad educativa a evaluar mientras se mantiene una perspectiva analítica basada en el propósito del trabajo de campo: evaluar.
h) Diferenciar claramente descripciones, interpretaciones (propias y ajenas) y juicios de valor (propios y ajenos).
i) Aportar *un feedback* formativo como una parte del proceso de verificación del trabajo de campo. Regular *el feedback* cuidadosamente. Observar su impacto.
j) En las notas de campo y en el informe de la evaluación incluir experiencias, pensamientos e impresiones propias. Esta información también constituye los datos de campo.

CAPÍTULO 10

El estudio de casos

El estudio de casos, como método de investigación, ha tenido gran importancia en el desarrollo de las ciencias sociales. La historia está llena de «casos» relevantes o «rupturas conceptuales» que han surgido del uso del estudio de casos. Así, estudios como los de Freud, Piaget, Maslow, Adler, Rogers, etc., que nacieron y se nutrieron del estudio de casos, son ejemplos de la profunda influencia que este método ha tenido en la ciencia.

El estudio de casos constituye uno de los modelos-tipo generales de investigación en las ciencias sociales (Goetz y LeCompte, 1988), siendo el enfoque tradicional de toda investigación clínica. Consiste en una descripción y análisis detallados de unidades sociales o entidades educativas únicas (Yin, 1989).

Ante los grandes enfoques metateóricos, el nomotético, que estudia muestras de población, y el idiográfico, que estudia un caso en profundidad, el estudio de casos se incluye dentro de este último, cuya finalidad se orienta a la compresión profunda de la realidad singular: individuo, familia, grupo, institución social o comunidad (De la Orden, 1985).

El estudio de casos es un diseño de investigación particularmente apropiado para estudiar un caso o situación con cierta intensidad en un período de tiempo corto (aunque hay casos que pueden durar más). La fuerza del estudio de casos radica en que permite concentrarse en un caso específico o situación e identificar los distintos procesos interactivos que lo conforman. Estos procesos pueden permanecer ocultos en un estudio de muestras (Walker, 1982).

Si bien el estudio de casos se describe aquí desde la perspectiva humanístico-interpretativa, y se pone más énfasis en su vertiente cualitativa, no quiere decir que no pueda contemplarse desde otras perspectivas con un enfoque más cuantitativo.

10.1 CONCEPTO Y CARACTERÍSTICAS

El estudio de casos es la forma más propia y característica de las investigaciones idiográficas llevadas a cabo desde una perspectiva cualitativa y debe considerarse como una estrategia encaminada a la toma de decisiones. Su verdadero poder radica en su capacidad para generar hipótesis y descubrimientos, en centrar su interés en un individuo, evento o institución, y en su flexibilidad y aplicabilidad a situaciones naturales.

Ha sido definido por Aldeman (1977) como un término paraguas que engloba una familia de métodos de investigación que tienen en común el enfocar la indagación sobre un caso o ejemplo, y por Bogdan y Bilken (1982) como un examen detallado de una situación, de un único sujeto, de un único depósito de documentos o de un evento particular.

Las notas que parecen traducir los criterios utilizados por los autores para definir los límites de esta categoría son: el estudio intensivo del caso y el empeño por llegar a la comprensión de la singularidad objeto de estudio. Merriam (1988, 11-13) señala cuatro propiedades esenciales del estudio de casos, las de ser particular, descriptivo, heurístico e inductivo. Es particular en cuanto que se centra en una situación particular, evento, programa o fenómeno; es descriptivo porque pretende realizar una rica e intensiva descripción del fenómeno estudiado; es heurístico en tanto que el estudio ilumina al lector sobre la comprensión del caso; y es inductivo, puesto que llega a las generalizaciones, conceptos o hipótesis partiendo de los datos.

El estudio de casos afronta la realidad mediante un análisis detallado de sus elementos y la interacción que se produce entre ellos y su contexto, para llegar mediante un proceso de síntesis a la búsqueda del significado y la toma de decisión sobre el caso. El estudio detallado permite clarificar relaciones, descubrir los procesos críticos subyacentes e identificar fenómenos comunes (Hamilton y Delamont, 1974).

El investigador puede estudiar el impacto de una innovación específica en una escuela, una experiencia de renovación curricular, o puede evaluar una institución escolar, un día en la vida de un profesor, de un alumno o de un adulto. Estos estudios tan diferentes tienen en común cierta dedicación al conocimiento y descripción de lo idiosincrático y específico como legítimo en sí mismo (WaIker, 1982).

Para Ary y otros (1987, 308) el estudio de casos puede plantearse objetivos como los siguientes:

a) Describir y analizar situaciones únicas, como pueden ser un niño superdotado, un nuevo modelo de dirección de centros.

b) Generar hipótesis que contrastar posteriormente en otros estudios más rigurosos. Los estudios de Piaget sobre la maduración de la inteligencia originaron hipótesis que posteriormente han sido investigadas con otros métodos.

e) Adquirir conocimientos.

d) Diagnosticar una situación, para orientar, llevar acabo un asesoramiento, recuperación, acción terapéutica, reeducación (ámbito clínico). Los estudios de Freud para ayudar a sus pacientes a resolver sus problemas de personalidad son ejemplos de este objetivo.

e) Completar la información aportada por investigaciones estrictamente cuantitativas.

La naturaleza del caso puede ser muy heterogénea (sujeto, grupo, institución, programa, etc.) y en parte condiciona el nivel descriptivo (tipo crónica, listado de rasgos, evaluación, intentos de contrastación), interpretativo, evaluativo o varios a la vez.

10.2 DISEÑO DEL ESTUDIO DE CASOS

Como en toda investigación educativa, en el estudio de casos se planifica, se recogen datos, se analiza e interpreta la información y se elabora el informe. El diseño del estudio de casos se inscribe dentro de la lógica que guía las sucesivas etapas de

recogida, análisis e interpretación de los datos de los modelos cualitativos, con la peculiaridad de que el propósito de la investigación es el estudio intensivo y profundo de uno o pocos casos de un fenómeno.

El diseño se articula en tomo a una serie de pasos o fases que siguen un enfoque progresivo e interactivo: el tema se va delimitando y focalizando a medida que el proceso avanza. Las primeras fases son de exploración y reconocimiento, se analizan los lugares, situaciones y sujetos que pueden ser materia o fuente de los datos, y las posibilidades que revisten para los fines y objetivos de la investigación. En la fase intermedia se seleccionan los sujetos o aspectos por explorar, las personas a entrevistar, qué estrategias se van a utilizar, la duración del estudio, etc.

Cubiertos estos pasos se pasa a la fase de recogida, análisis e interpretación de los datos, para terminar con la elaboración del informe y la toma de decisiones.

El investigador, a medida que va cubriendo las fases del estudio, incorpora las nuevas ideas y planteamientos que van surgiendo, lo que permite modificar o reestructurar las anteriores; este procedimiento es recurrente a 10 largo de todo el tiempo que dura el estudio.

10.3 ESTRATEGIAS E INSTRUMENTOS DE RECOGIDA DE DATOS

Aunque la observación y entrevista suelen ser utilizadas con frecuencia en el estudio de casos, se incluye cualquier técnica que el diseño de investigación precise. Los métodos e instrumentos de recogida de datos se seleccionan por su propiedad para la investigación.

Los datos del estudio de casos suelen proceder de diversas fuentes: de entrevistas, del estudio de documentos personales, como diarios, cartas; de evaluaciones psicológicas, físicas o sociológicas; de interrogar a otras personas, como padres, familiares o amigos; del análisis de documentos; etc.

Patton (1978) considera que los estudios de casos incluyen datos procedentes de la observación, las entrevistas, documentales, terceras personas, y todos ellos en su seguimiento a lo largo del tiempo, con 10 que llegará a almacenarse gran cantidad de información.

IOA TIPOS DE ESTUDIOS DE CASOS

- *Estudio de casos institucional.* Se centra en el estudio de una organización particular durante un tiempo, describiendo y analizando desarrollo de la misma. El estudio puede versar sobre la historia de una escuela, cómo surgió, cómo fueron sus primeros años, cambios que han ocurrido, cómo es ahora, etc.
- *Estudio de casos observacional.* En estos estudios la técnica de recogida de datos es la observación participante y el foco de estudio una, determinada organización (escuela, centro de reeducación), algún aspecto de la institución (un aula, una sala de profesores o un grupo específico de profesores), alguna actividad de la escuela (estudio de una innovación, del currículum), o combinación de estos aspectos.
- *Las historias de vida.* Es una forma de estudio de casos. El investigador realiza entrevistas con alguna intensidad a una persona con el propósito de recoger información

de primera mano. Cuando es hecha por historiadores se refiere a la historia oral. La factibilidad de estos estudios de casos viene determinada por la naturaleza del sujeto potencial.

- *Los estudios comunitarios.* Son similares a los estudios organizacionales y observacionales, con la excepción de que el foco de interés es la comunidad o vecindad.
- *El análisis situacional.* Este tipo de estudios de casos se centra en un evento particular (la expulsión de un alumno de la escuela, por ejemplo) y lo estudia desde diversos puntos de vista: alumno, profesor, padres, compañeros, director.
- *La microetnografía.* En el estudio de casos se refiere al estudio de una unidad muy pequeña de una organización, no de una actividad (niños aprendiendo a dibujar).
- *Los estudios de casos múltiples.* Cuando los investigadores estudian dos o más sujetos, situaciones o depósitos de datos, están haciendo estudios de casos múltiples. Hay diseños de investigación diferentes de los señalados anteriormente que exigen el estudio de casos múltiples. La lógica que emplean es diferente de la de los enfoques de estudios de casos porqu~ se orientan a desarrollar teoría y requieren más de dos o tres situaciones o sujetos.

10.5 DIFICULTADES QUE PRESENTA EL ESTUDIO DE CASOS

El estudio de casos como modalidad de investigación presenta diversas dificultades metodológicas, algunas específicas y otras compartidas con otros diseños de investigación. Una de las dificultades estriba en cómo formular generalizaciones a partir de una realidad singular, que queda enmarcada en dos preguntas: ¿cómo puede justificar algo el estudio de un solo caso?, y aunque fuese justificable teóricamente, ¿qué empleo puede hacerse del estudio de casos por parte de quienes han de adoptar una acción? (Walker, 1983).

Para Campbell y Stanley (1963) el estudio de casos con una sola medición constituye el tipo de diseño más vulnerable a las fuentes de invalidez, tanto interna como externa. La crítica al estudio de casos pone el acento en este y otros problemas. Señalan el hecho de que la generalización no es posible y cuestionan el valor de los estudios de un solo evento (Bell, 1987,6-7).

Pese a su vigencia en el campo de la investigación, su rigor científico presenta algunas dificultades. Kratochwill (1977) las resume en las siguientes: *a)* falta de atención a la validez, tanto interna como externa; *b)* sus limitadas opciones de diseño; c) la dificultad existente para generalizar los hallazgos.

mayor exhaustividad el diseño respondente por ser el que refleja mejor los elementos esenciales de cualquier diseño evaluativo.

1) Diseño iluminativo

Este diseño (Parlett y Hamilton, 1977), más que centrarse sólo en la valoración del producto educativo, enfatiza el estudio intensivo del programa como *totalidad:* sus principios básicos, su evolución, sus actividades, sus logros y dificultades. Básicamente evalúan el sistema de *enseñanza* y el *medio de aprendizaje.* El *sistema de enseñanza* está constituido por planes de estudio, programas, supuestos pedagógicos, detalles acerca de las técnicas y del material. Dado que al aplicar un sistema de enseñanza (modelo abstracto) aparecerán notables modificaciones según el contexto, los alumnos y el profesor, la evaluación habrá de llevarse a cabo inmersa en las condiciones concretas donde se aplica el sistema.

El *medio de aprendizaje* se refiere al contexto material, psicológico y social en el que trabajan conjuntamente profesores y alumnos. Quedan implicadas variables culturales, sociales, institucionales y psicológicas que interactúan originando un único estado de circunstancias (presiones, opiniones, conflictos) que impregnan los procesos de enseñanza y aprendizaje.

Los problemas que hay que investigar suelen centrarse en el análisis de los procesos interactivos que tienen lugar en el aula. Su constante dinamismo origina una realidad cambiante, cuya evaluación exige la implicación y familiarización intensiva por parte del evaluador. Los problemas no pueden definirse previamente, aparecerán progresivamente durante el transcurso de los acontecimientos.

El proceso global comprendería etapas generales como las siguientes:

a) Etapa de exploración. Se observa la mayor gama posible de variables que pueden afectar al resultado del programa o de la innovación. Es una fase de reconocimiento.

b) Etapa de investigación. Se seleccionan y plantean cuestiones con el fin de elaborar una lista sistemática y selectiva de los aspectos más importantes del programa en su contexto.

c) *Etapa de explicación.* Los principios generales subyacentes a la organización del programa son especificados y se delinean posibles relaciones de causalidad.

2) Diseño democrático

Stenhouse (1984, 1987), McDonald (1976) y Elliott (1978) son considerados por Pérez Gómez *(1985b)* como los principales representantes de un modelo de evaluación que pretende desarrollar al máximo las posibilidades de los nuevos enfoques cualitativos. Estos autores ponen de relieve la necesidad de investigar y evaluar la vida del aula con planteamientos y procedimientos metodológicos naturalistas. Si la evaluación ha de servir para la realidad educativa, esta habrá de ser analizada en el marco de la realidad diaria.

La investigación trata de orientar la actividad educativa ayudando a decidir *qué* hay que hacer y cómo hacerlo. El evaluador ha de sumergirse en el contexto real y vital en que tienen lugar los acontecimientos y conocer las interpretaciones diversas que hacen los participantes que las viven directamente. De ahí que, tanto las fuentes de datos como los destinatarios de los informes, son todos los que participan en el programa. Por esta razón se denomina *diseño democrático de evaluación.* El propósito básico es facilitar y promover el cambio, llegando a una transformación real mediante la modificación de concepciones , creencias y modos de interpretar de las personas implicadas.

C) DISEÑO RESPONDENTE

Para Stake (1967, 1975 Y 1976), la evaluación es un valor observado y comparado con alguna norma y ha de enfatizar lo que se ha observado acerca del programa, detectando la satisfacción o insatisfacción que muestran las personas implicadas en el programa. La evaluación respondente contempla exhaustivamente aspectos como antecedentes, procesos o actividades del programa, sus efectos, datos para los juicios valorativos, el informe holístico y el asesoramiento a los educadores.

Teniendo en cuenta la creciente implicación de los docentes en los procesos de planificación, implantación y evaluación, el diseño *respondente*, por su carácter global, clarificador y didáctico, brinda un marco adecuado para que puedan producirse cambios reales en el ámbito educativo. En consecuencia, este diseño será objeto de una descripción más detallada.

11.5 DISEÑO RESPONDENTE

Stake (1975) elaboró un diseño centrado en responder a los problemas y cuestiones reales que se plantean las personas implicadas en un programa educativo. Según este diseño, el evaluador responde a la demanda presentada por diferentes audiencias, negociando con las personas implicadas lo que debe hacerse.

A pesar de las críticas dirigidas contra la evaluación orientada a objetivos, numerosos investigadores han seguido utilizando los objetivos como criterio de referencia, pero Stake (1967, 1975 Y 1983) introduce valiosas innovaciones. Para este autor, hay que tener en cuenta la justificación teórica o base lógica del programa que se va a elaborar, una matriz de descripción y otra matriz de juicio o valoración (fig. 11.6). El diseño exige completar estas dos matrices de información, cada una de las cuales está dividida, a su vez, en dos columnas más.

La primera columna (casillas 1, 2 Y 3) se destina a la descripción de la *planificación* del programa, y en ella se detallan los objetivos o intenciones. Esta información constituye el *modelo teórico* que se ha planificado. La segunda columna (casillas 4, 5 Y 6) se reserva para la descripción de lo que ocurre en la realidad durante la *implantación* o aplicación del programa tal como se ha planificado en la columna anterior. Esta información procedente de lo que ocurre en la realidad origina el *modelo empírico.* Las otras dos columnas abarcan los criterios o normas de calidad absolutas y relativas (casillas 7,8 Y 9) Y las valoraciones emitidas (casillas 10, 11 Y 12). Tanto

la matriz descriptiva como la de juicio contemplan antecedentes o características de la situación inicial, procesos o actividades, y efectos o resultados del programa.

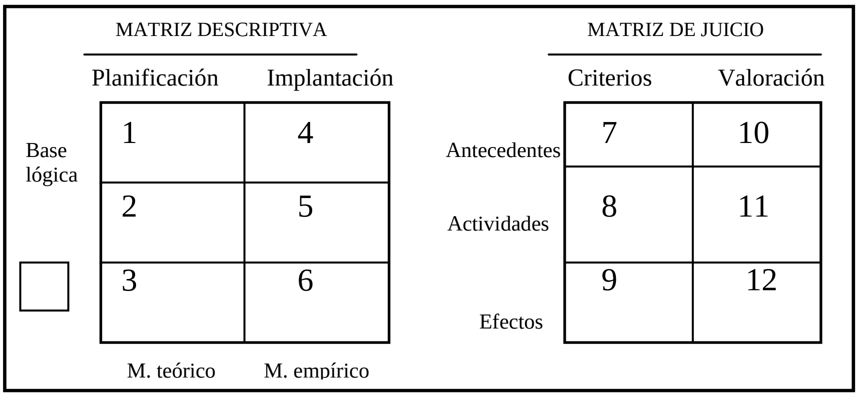


FIG. 11.6 Diseño respondente. Adaptado de Stake (1967)

Los *antecedentes* son condiciones existentes antes de que se inicie el proceso educativo o la implantación del programa y que pueden influir en los resultados. Por ejemplo, las características de las situaciones vitales que han de afrontar los sujetos que participarán en el programa. Estas características contextuales son muy cambiantes y exigirán actividades específicas, coherentes y adaptadas a dichas situaciones vitales. También puede ser necesario incluir aquí características descriptivas de los sujetos como inteligencia o rendimiento previo.

Las *actividades o procesos* didáctico s constituyen las actividades e interacciones de los alumnos con profesores, padres, tutores, compañeros y otros adultos. En general, los procesos representan el conjunto de interacciones y actividades que tienen lugar en el escenario educativo. Estas actividades planificadas constituyen el programa.

Bajo la denominación de *efectos* se incluye todo lo que se puede conseguir a través de un programa. Aquí están implicados tanto los efectos esperados como los no esperados. Los resultados son los efectos del proceso educativo y se refieren a la consecución de objetivos en relación a conductas terminales de los alumnos, a las condiciones contextuales y al comportamiento del profesor.

JUSTIFICACIÓN TEÓRICA DEL PROGRAMA

Está constituida por los planteamientos teóricos e intenciones filosóficas en que se basará el programa y sobre las que se planifica la estructura organizativa de una institución y, más en concreto, su proyecto educativo. Para enjuiciar el programa será imprescindible esta información, de ahí que tenga que ser elaborada cuidadosamente. Durante la elaboración del programa el evaluador ha de poder preguntarse: « ¿Constituye el programa que se planifica -

y se implanta un paso lógico en la realización de las intenciones básicas?».

El contenido de la *base lógica* permite analizar si la dinámica real durante la planificación e implantación del programa es coherente con las directrices y las intenciones básicas. A su vez, la *base lógica* sirve de pauta para la formulación de los objetivos del programa y también es útil para elegir grupos de referencia a los que se recurrirá para identificar los criterios de calidad y la valoración consecuente.

PLANIFICACIÓN DEL PROGRAMA

La planificación incluye los objetivos o intenciones del programa y describe todo lo que se planea, tomando conciencia de los aspectos u objetivos que se excluyen y de las razones que hay para no tenerlos en cuenta. Esta información configura el modelo teórico o conceptualización del programa y se desglosa en tres casillas temporalmente ordenadas tal como indica la numeración correspondiente a la primera columna y según los tres ámbitos ya mencionados: antecedentes o situación inicial, actividades y efectos deseados.

Casilla 1. *Previsión de la situación inicial.* Se describen los antecedentes o características previstas de los sujetos que se beneficiarán del programa y de otros que tendrán algún tipo de implicación en el mismo. También es necesario describir las características de la situación o contexto que constituirá el escenario previsto para la futura implantación de las actividades y procesos que se planifican en la casilla 2. En general, deben incluirse todas aquellas características relevantes que puedan arrojar luz sobre cuáles serán las actividades más adecuadas y que puedan condicionar los efectos esperados de la casilla 3. A modo de sugerencia puede resultar orientativo responder a preguntas como las formuladas en la figura 11.7.

Si se tuviera previsto utilizar un diseño experimental o cuasiexperimental para evaluar la efectividad de las actividades previstas, la información de la casilla 1 correspondería a la que se incluiría en el proyecto del experimento al planificar la situación pretest. En definitiva, se trata de prever cómo será la situación inicial a la que se destinará el programa. Por ejemplo, podemos planificar un programa de intervención para niños de primaria con un nivel sociocultural bajo, que estén sometidos a inmersión lingüística, obtengan puntuaciones inferiores al nivel medio de la clase en la prueba de Galí/Burt y dominen la mecánica de cálculo, pero no la comprensión del valor posicional de las cifras.

¿Qué se hace?	*¿Porqué se hace?*	*¿Quién lo hace?*	*¿Para quién?*	*¿Para qué?*	*¿Cuándo?*	*¿Dónde?*	*¿Cómo?*	*¿Con qué?*	*¿Resul tados?*

FIG. 11.7 Sugerencias para describir las características previstas para la situación inicial donde se aplicará el programa

Casilla 2. *Actividades previstas.* Dadas las características que se habrán descrito en la fase anterior, ahora se planifica cómo ha de ser el proceso educativo, qué se hará, es decir, en qué consistirá la intervención o tratamiento educativo para obtener los resultados que se incluirán en la casilla 3. Como puede apreciarse, la información de la casilla 2 correspondería a la que se incluiría en el proyecto de **un** experimento al planificar el tratamiento que se aplicará o variable independiente.

De forma orientativa, el evaluador puede formularse cuestiones como las que aparecen en la figura 11.8.

¿Qué debe hacerse?	*¿Qué puede hacerse?*	*¿Quién lo hará?*	*¿Para quién?*	*¿Para qué?*	*¿Cuándo?*	*¿Dónde?*	*¿Cómo?*	*¿Co n qué?*

FIG. 11.8 Posibles cuestiones que orientan la planificación del programa

Al plantearse cuestiones como las de la figura 11.8 pueden confeccionarse esquemas como el que se reproduce en la figura 11.9 para evitar que se omitan aspectos importantes.

Por ejemplo, con el fin de mejorar el cálculo mental, el profesorado de **un** centro de primaria puede introducir juegos con dados y fichas de colores, agrupando los alumnos en grupos de cinco durante tres días a la semana. En cada casilla de la matriz de la figura 11.1 puede incluirse información relacionada con la secuenciación de las actividades.

Profesor	Grupo	Actividad	Material	Agrupación sujetos	Frecuencia duración	Objetivo

FIG. 11.9 Sugerencias para describir la planificación del programa

Así, durante la primera y segunda semanas se realizan juegos con fichas de madera coloreadas; durante la segunda semana se introducen además juegos con dados hasta la cuarta, mientras que las fichas se suprimen al acabar la segunda semana.

		Temporalización				
		1.ª semana	2.ª semana	3.ª semana	4.ª semana	etc
Actividades planificadas	juego fichas	xxxxxx	xxxxxx			
	juego datos		xxxxxx	xxxxxxxx	xxxxxxxx	
	etc.					

FIG. 11.10. Extracto de una matriz con la temporalización de actividades y el período respectivo de implantación (X)

Casilla 3. *Efectos esperados.* Por último, los resultados esperados se incluyen en la casilla 3. Esta información correspondería a la que se incluiría en el proyecto del experimento al planificar la situación postest para medir el nivel de la variable dependiente después de que se aplique la variable independiente.

Así, al introducir juegos con fichas de madera coloreadas podemos esperar que los alumnos comprendan mejor el valor posicional de las cifras, lo que favorece que los alumnos asimilen recursos que les permitan agilizar el cálculo mental a través de juegos con dados. En la figura 11.11 se registran horizontalmente los tipos o niveles de efectos esperados y verticalmente las actividades que se han planificado con el fin de alcanzar dichos resultados. Sin embargo, generalmente, no se espera que las distintas actividades contribuyan por igual al logro de cada efecto previsto, de ahí la conveniencia de incluir símbolos (+ -) en la matriz para indicar una estimación de la contribución relativa de cada actividad a los distintos efectos. También es muy sugerente anotar en las casillas de la misma matriz o de otra preparada al efecto los comentarios de las personas implicadas o de expertos en cada ámbito, en torno a las expectativas que mantienen acerca de la posible contribución de cada actividad para lograr cada efecto propuesto.

PARTE IV

INVESTIGACIÓN MÁS ORIENTADA A LA PRÁCTICA EDUCA TIV A: DECISIÓN Y CAMBIO

La investigación orientada a la *práctica educativa* es la que se diseña y realiza con el propósito de proporcionar información sobre problemas prácticos, para tomar decisiones, evaluando la implantación de una determinada política o estimando los efectos de la política existente. Por tanto, a diferencia de las *perspectivas empírico-analítica y humanístico-interpretativa,* la finalidad esencial de la investigación orientada a la *práctica educativa* no es tanto acumular conocimientos sobre el proceso educativo y explicar y/o comprender la realidad educativa como aportar información que guíe la toma de decisiones y los procesos de cambio para mejorar la práctica educativa.

Dado el contexto aplicado de esta perspectiva, la investigación orientada a la *práctica educativa* no suele poseer una metodología con entidad propia y, en consecuencia, puede utilizar las dos grandes vías metodológicas que predominan, respectivamente, en las orientaciones *empírico-analítica y humanístico interpretativa* (De Miguel, 1988, y Nisbet, 1988).

Dentro de esta orientación destacan la *investigación evaluativa,* decisiva para la toma de decisiones y orientada a determinar la eficacia de organizaciones y programas educativos, y la *investigación acción,* más orientada a promover e implantar el cambio, a mejorar la capacidad de autorreflexión, a guiar la elaboración del currículo y potenciar la formación del propio educador, de los estudiantes y de las demás personas implicadas.

La investigación *participativa*, orientada al cambio social y al desarrollo comunitario, y la investigación cooperativa o colaborativa, por su mayor énfasis en la cooperación de grupos de participantes, pueden considerarse vinculadas a la *investigación acción.*

CAPÍTULO **11**

Investigación evaluativa

Durante la presente década la *investigación evaluativa* ha conocido un desarrollo sin precedentes (Cordray y Lipsey, 1987b, y Keeves, 1988). Ello ha sido debido, en parte, a la creciente necesidad de constatar la eficacia relativa de los múltiples .programas educativos y otras alternativas innovadoras en cuya planificación y aplicación se invierten considerables sumas de dinero, numerosos planes de investigación y renovados esfuerzos en tomo a la formación del profesorado.

11.1 CONCEPTO Y CARACTERÍSTICAS

El concepto de *investigación evaluativa* es muy amplio y su definición depende de las concepciones previas de *evaluación* y de *programa* (De la Orden, 1985; Cordray y Lipsey, 1987b; Stufflebeam y Shinkfield, 1987).

En general, la *evaluación* puede considerarse como un proceso, o conjunto de procesos, para la obtención y análisis de información significativa en que apoyar *Juicios de valor* sobre un objeto, fenómeno, proceso o acontecimiento, como soporte de una eventual decisión sobre el mismo. Esta *decisión* dependerá, en parte, del grado de adecuación a un *elemento referencial o criterio.*

La denominación genérica de *programa* suele referirse a cursos de acción de carácter sistemático que se inician con la intención de alcanzar determinados objetivos.

En consecuencia, *la evaluación de programas* puede definirse, de un modo general, como el proceso sistemático de recogida y análisis de información fiable y válida para tomar decisiones sobre un programa educativo (De la Orden, 1985). En la medida que este proceso sea riguroso, controlado y sistemático, constituye un modo de investigación que hoy se conoce con el nombre de *investigación evaluativa* (Scriven, 1986, y Herman *et al.*, 1987). Se trata, pues, de una forma de investigación educativa aplicada que intenta determinar el valor de programas educativos de acuerdo con determinados criterios. Para ello se relacionan procesos y efectos y, en base a la información obtenida, se valoran para tomar decisiones presumiblemente optimizantes de la situación. De ahí que la *investigación evaluativa* se caracterice no sólo por la

proyección práctica inherente a toda investigación aplicada (Nisbet, 1988, y Keeves, 1988), sino, esencialmente, por ser un proceso presidido *por juicios de valor* (Scriven, 1987). La *investigación evaluativa* se orienta a valorar una situación concreta (programa) y tomar decisiones alternativas; sin embargo, la *investigación básica* apunta a la producción de la teoría, o explicación de fenómenos mediante la determinación de relaciones entre variables.

11.2 PROCESO DE INVESTIGACIÓN

Las características de la *investigación evaluativa* sugieren que esta modalidad de_ investigación puede constituir una vía decisiva para el progreso teórico y para la optimización de la praxis educativa (Stuftlebeam y Shinkfield, 1987).

El proceso general de investigación, que sirve de punto de referencia para desarrollar cada modalidad de investigación, presenta algunas especificaciones en el caso de la *investigación evaluativa.* Así, De la Orden (1985, 134-135), comparándola con la investigación estricta, caracteriza la *investigación evaluativa* con los siguientes matices diferenciadores:

a) Los juicios de valor se explicitan no sólo en la selección y definición del problema, cuya responsabilidad comparte el evaluador con otras instancias, aquellas de las que el programa depende, sino también en el desarrollo y aplicación de los procedimientos del estudio.
b) Resulta difícil, y a veces inadecuado, formular hipótesis precisas.
e) La replicación es prácticamente imposible, dado que el estudio de evaluación está íntimamente vinculado a un programa determinado en una situación específica y en un momento concreto de su desarrollo.
f) Los datos que hay que recoger están condicionados por la viabilidad del proceso. Las opciones, cuando son posibles, reflejan juicios de valor de los administradores dores de los programas. Con frecuencia se constatan diferencias entre datos accesibles y datos deseables para el evaluador y para los responsables del programa.
e) Muchas variables relevantes sólo superficialmente son susceptibles ge control. La aleatorización, como técnica general de control de la varianza sistemática debida a la influencia de variables extrañas al estudio, es muy difícil de conseguir en la investigación evaluativa.
1) El criterio de decisión, en orden a la continuación, modificación, ampliación o sustitución del programa, es responsabilidad de la agencia administradora del programa o de la audiencia a quien se dirige el informe de evaluación.
g) El informe evaluativo debe adaptarse a las exigencias de quien toma las decisiones sobre el programa.

Como sugiere De la Orden (1985, 135), a pesar de estas diferencias, el diseño de la investigación evaluativa no presenta variaciones sustanciales con respecto al proceso general de investigación (fig. 11.1). Es más, aunque la investigación evaluativa puede basarse en diseños genuinamente evaluativos, cualquiera de los métodos y diseños abordados en el marco de las perspectivas *empírico-analítica y humanístico-interpretativa* pueden resultar también potencialmente adecuados (Fitz-Gibbon y Morris, 1987b).

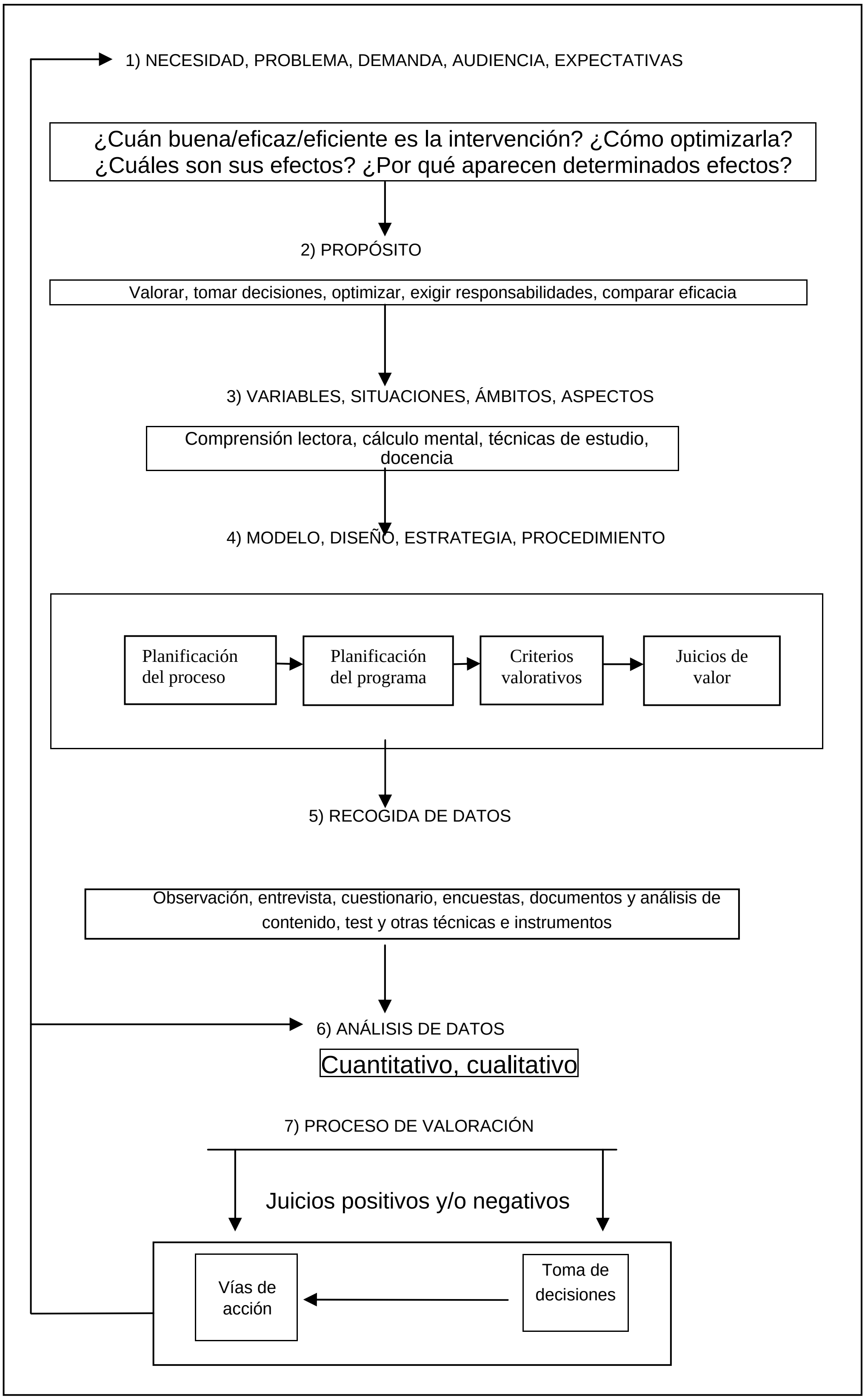


FIG. 11.1 Ejemplificación de posibles fases de la investigación evaluativa

Numerosas situaciones problemáticas en el ámbito educativo requieren la realización de una investigación evaluativa. Así, algunas decisiones valorativas asociadas a la investigación evaluativa están implicadas en cuestiones como las siguientes:

- ¿Qué programa de lectura es más eficaz?
- ¿Cuáles son los factores que caracterizan un programa de eficiencia lectora?
- ¿Cómo progresan los alumnos sometidos a un programa de técnicas de estudio?
- ¿Cuál es la naturaleza de las interacciones entre profesores y alumnos al aplicar un programa de formación en la empresa?
- ¿Cómo se realiza la aplicación del programa? ¿Se hace con arreglo al plan previsto?
- ¿Qué vivencias tienen los alumnos? ¿Qué actividades realizan? ¿Cómo actúan los profesores?
- ¿Cómo está organizado el programa? ¿Cómo es el contexto en el que se explica?
- ¿Es eficaz el programa de formación propuesto?
- ¿Se logran mejor los objetivos con el método analítico que con el método sintético?
- ¿Para qué tipos de alumnos es adecuada la estrategia basada en técnicas de simulación?
- ¿Cuáles son los aspectos más eficaces de un programa de formación?

Aunque la *investigación evaluativa* puede dirigirse a determinar el valor de un programa en acción (Fetterman y Pitman, 1986), tratando de contestar a las preguntas acerca del logro de sus objetivos o de su correcto funcionamiento (King *et al., 1987),* más frecuentemente se aplica a nuevos proyectos educativos para constatar su presunta superioridad sobre otros programas ya existentes (De la Orden, 1985).

11.3 CONSIDERACIONES METODOLÓGICAS

La *investigación evaluativa* constituye una síntesis, al mismo tiempo que una extensión de una amplia gama de métodos de investigación que aportan información sobre cuestiones planteadas en tomo a los programas educativos, con el fin de facilitar la toma de decisiones sobre los mismos.

Bajo la denominación de *paradigma* hemos englobado metodologías compartidas por investigadores y educadores que adoptan una determinada concepción del proceso educativo (Escudero, 1987, y De Miguel, 1988). Cada paradigma se caracterizaba por una forma común de investigar en el seno de una comunidad científica y hemos considerado las perspectivas *empírico-analítica* y *humanístico-interpretativa.* Somos conscientes de que la teoría crítica ha originado un paradigma recientemente utilizado en la investigación educativa; sin embargo, desde la óptica metodológica que nos ocupa, y según De Miguel (1988) y Keeves (1988), el paradigma *crítico* participa de los postulados que caracterizan al paradigma interpretativo, añadiendo un componente ideológico.

Cada perspectiva presentaba ventajas e inconvenientes. Aunque parten de supuestos diferentes, era posible conjugar las aportaciones de ambas perspectivas metodológicas desde una postura *ecléctica.* Las características del fenómeno educativo que hay que evaluar y los elementos contextuales constituyen criterios para delimitar hasta dónde podemos llegar al utilizar las respectivas metodologías en cada situación educativa. De ahí que los paradigmas no sean considerados como determinantes únicos de la metodología que hay que emplear al evaluar los programas. El nexo

existente entre paradigma y método orienta la elección de la metodología, pero no la determina de forma inexorable.

Reichardt y Cook (1986) señalan que la perspectiva paradigmática del investigador ha de ser flexible y capaz de adaptaciones. Algunos ámbitos de la evaluación de programas reclaman una creciente utilización combinada de las metodologías asociadas a cada paradigma En el ámbito de la investigación evaluativa es probable que convenga conceder una mayor relevancia a los métodos cualitativos por centrarse no tanto en la verificación y confirmación, más propias de la metodología cuantitativa, como en el descubrimiento. Aunque es de esperar que la utilización óptima de las posibilidades de cada metodología se dé en el marco de su paradigma respectivo, es adecuado beneficiarse de lo mejor de ambos a través de su integración. Puede concluirse, pues, que ambas modalidades de investigación son distintas pero no incompatibles, aunque partan de supuestos epistemológicos distintos.

11.4 DISEÑOS DE INVESTIGACIÓN EV ALUA TIV A

La clasificación de los diseños evaluativos, considerados como marcos generales de referencia para planificar una investigación evaluativa, ha generado una abundante literatura. A pesar de que no existe un consenso en las distintas clasificaciones, las aportaciones de autores como Parton (1980), Guba y Lincoln (1982), Pérez Gómez (1983), y Stufflebeam y Shinkfield (1987), presentan entre sí muchas afinidades (tabla 11.1), y en ellas, de manera más o menos explícita, subyace el criterio de la diversidad y complementariedad metodológica.

TABLA 11.1 Clasificaciones de los diseños según distintos autores

Perspectiva	*Patton* (1980)	*Guba-Lincoln* (1982)	*Pérez Gómez* (1983)	*Stufflebeam Shinkfield (1987)*	*Autor*
Empírico-analítica	Objetivos	Objetivos	Objetivos	Objetivos	Tyler (1950)
	Análisis sistemas		Análisis sistemas		Rivlin (1971) Rossi (1979)
				Método científico	Suchman (1967)
Humanístico-interpretativa	Respondente	Respondente	Respondente	Respondente	Stake (1975)
	Iluminativo		Iluminativo	Iluminativo	Parlett y Hamilton (1977)
	Sin metas	Sin metas		Sin metas	Scriven (1967)
			Democrático		McDonald(1976)
Susceptibles de comple-mentariedad	UTOS		UTOS		Cronbach (1963)
	CIPP	CIPP	CIPP		Stufflebeam (1966)
	Crítica artística	Crítica artística	Crítica artística		Eisner (1971)
	Adversario			Contrapuesto	Wolf (1974)

Describiremos los diseños que tienen una orientación *empírico-analítica,* los que son más susceptibles de complementariedad metodológica y los que siguen la perspectiva humanístico-interpretativa, tomando como referencia directa las aportaciones de autores como Pérez Gómez (1983), Fetterman y Pitman (1986), Cordray y Lipsey *(1987a),* Guba y Lincoln (1982 y 1987) Y Stufflebeam y Shinkfield (1987), teniendo presente que, en la práctica, casi nunca se aplican diseños puros y más que metodologías diferenciadas conviene considerar un continuo metodológico (fig. 11.2).

Objetivos
Análisis de sistemas
Método científico

Respondente
Iluminativo
Sin metas
Democrático

PERSPECTIVA EMPÍRICO-ANALÍTICA

CONTINUO METODOLÓGICO

PERSPECTIVA HUMANÍSTICO-INTERPRETATIVA

UTOS
CIPP
Contrapuesto

FIG. 11.2 Diseños de evaluación situados en el continuo metodológico

DISEÑOS DE ORIENTACIÓN EMPÍRICO-ANALÍTICA

La evaluación basada en *objetivos* de comportamiento (Tyler, 1950) y en el análisis de sistemas (Rivlin, 1971), así como los diseños que siguen el modelo del *método científico* en su acepción positivista (Suchman, 1967), abordan variables y resultados previamente especificados y medidos de forma cuantitativa, recurriendo en la medida de lo posible a la manipulación experimental (Weiss, 1975, y Fitz-Gibbon y Morris, 1987 *b).* El hecho de considerarlos como diseños de orientación *cuantitativa* no supone que nunca utilicen técnicas y datos *cualitativos.* Incluso podría recurrirse a la *triangulación* recogiendo ambos tipos de datos en una misma investigación. Por su mayor relevancia, ofrecemos una breve descripción de los diseños basados en *objetivos* y en el *método científico.*

A) DISEÑO BASADO EN OBJETIVOS

Según el diseño de Tyler (1950), las decisiones acerca de los programas deben basarse en la coincidencia o congruencia entre los objetivos del programa y sus resultados. De ahí que las técnicas utilizadas consistan esencialmente en recopilar y analizar datos relativos a los *objetivos específicos.* Teniendo en cuenta que el objetivo de un programa es producir un cambio en el comportamiento del alumno, es necesario

medir el cambio producido en aspectos vinculados al programa. Ahora bien, en la práctica, tiende a medirse el cambio sólo en términos de logros del alumno y suele concebirse la evaluación como un proceso *Terminal* y centrado en objetivos previamente definidos. En este diseño pueden destacarse fases orientativas como las siguientes:

1. Identificar los objetivos de comportamiento del programa.
2. Seleccionar o elaborar instrumentos que permitan medir y valorar el logro de los estudiantes.
3. Aplicar estos instrumentos después de implantar el programa.
4. Comparar objetivos del programa y logros de los alumnos.

B) DISEÑO BASADO EN EL MÉTODO CIENTÍFICO

Este diseño se ciñe, en la medida de lo posible, a los supuestos del *método científico* con el fin de constatar relaciones causales entre elementos del programa y los resultados. Según Suchman (1967), la evaluación debe basarse en la lógica del *método científico* o, al menos, en alguna variación o adaptación de este modelo. Aunque el diseño ideal es el *experimental,* en el ámbito educativo se busca el equilibrio entre el rigor del método científico y la situación específica en la que se desarrolla la investigación evaluativa.

Para Suchman el propósito de la investigación evaluativa es también determinar hasta qué punto un programa consigue el resultado deseado. Los elementos fundamentales que intervienen en el diseño son (fig. 11.3): valores, objetivos, criterios, programa, implantación y efectos.

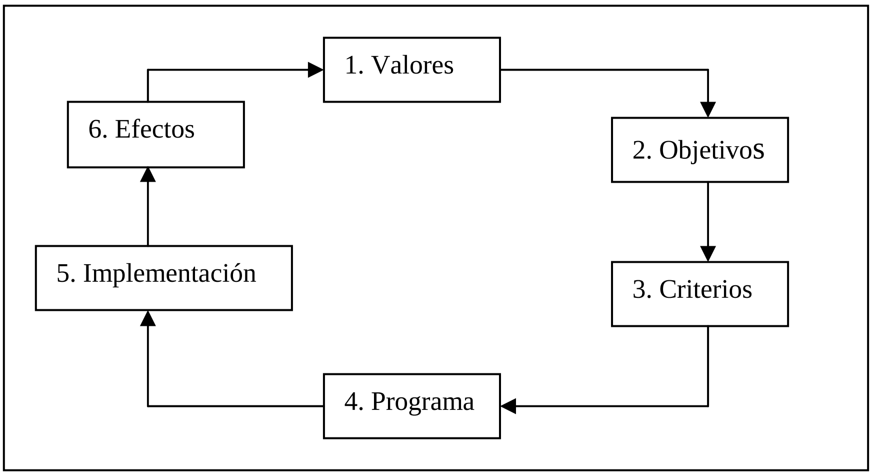


FIG. 11.3 Elementos fundamentales del diseño (adaptado de Suchman, 1967)

1) *Valores.* El diseño se inicia con procesos de valoración y culmina también con la valoración de los efectos del programa. Se entiende el valor como la cualidad de cualquier aspecto de una situación, actividad u objeto que puede presentar, en menor o mayor grado, características como utilidad, interés, adecuación o deseabilidad.

2) Objetivos. Los valores determinan los objetivos que habrá de cubrir el programa y orientan la elaboración de los criterios y la planificación del programa. Supongamos que un grupo de educadores considera valioso optimizar su práctica educativa. A partir de esta valoración puede proceder se a formular un objetivo como puede ser capacitar a dichos educadores para realizar investigaciones en su aula.

3) Criterios. Permiten valorar la consecución de los objetivos. Son criterios de éxito o de fracaso para juzgar el programa. Algunos criterios en el ejemplo planteado pueden ser:

- *Esfuerzo.* Calidad y cantidad de actividades planificadas en el programa de investigación educativa. Este criterio permitiría valorar las actividades del programa y no sus efectos.
- *Logros.* Resultados del esfuerzo. Por ejemplo, capacidad de los educadores para hacer un estudio crítico de determinados artículos y elaborar proyectos de investigación.
- *Suficiencia de los logros.* Grado en que la preparación que proporciona el programa cubre las necesidades de los educadores para que puedan investigar. Por ejemplo, el resultado suele valorarse según criterios como el potencial de utilidad.
- *Eficiencia.* Adecuación entre los resultados obtenidos y el esfuerzo realizado.
- *Proceso.* Determinación de causas que explican los resultados, es decir, características del programa, participantes implicados en el programa, contexto de aplicación y efectos del programa.

4) Programa. Planificar actividades para alcanzar los objetivos. Pueden formularse una serie de hipótesis que afirman que las actividades A, B y C, que configuran la variable independiente (VI), producirán los resultados X, Y y Z o variables dependientes (VD). Algunas actividades en el ejemplo propuesto pueden ser leer e interpretar críticamente informes de investigación, elaborar proyectos de investigación, organizar debates en tomo a varios artículos, plantear y resolver problemas prácticos.
También se tienen en cuenta las precondiciones (a_1, a_2, a_3,… a_n) y otros acontecimientos (b_l, b_2, b_3, … b_n) que pueden influir en la eficacia del programa inicial (fig. 11.4). Así, la experiencia previa de los educadores en investigación educativa puede matizar los efectos del programa, y el horario en que se aplica el programa y la actuación del profesor que lo imparte constituyen acontecimientos decisivos que también pueden influir en los resultados.

5) Implantación. Aplicación del programa. Se realizan las actividades con arreglo a la planificación prevista.

6) Efectos. Se recoge información sobre el nivel de la variable dependiente (VD), considerada como efecto de la aplicación del programa. Se contemplan efectos deseados e indeseados, reconocibles gracias a determinadas consecuencias (c_l, c_2, c_3, ... c_n), como pueden ser las puntuaciones de una prueba de conocimientos sobre investigación o los cambios introducidos en la dinámica de la clase impartida por cada educador participante.

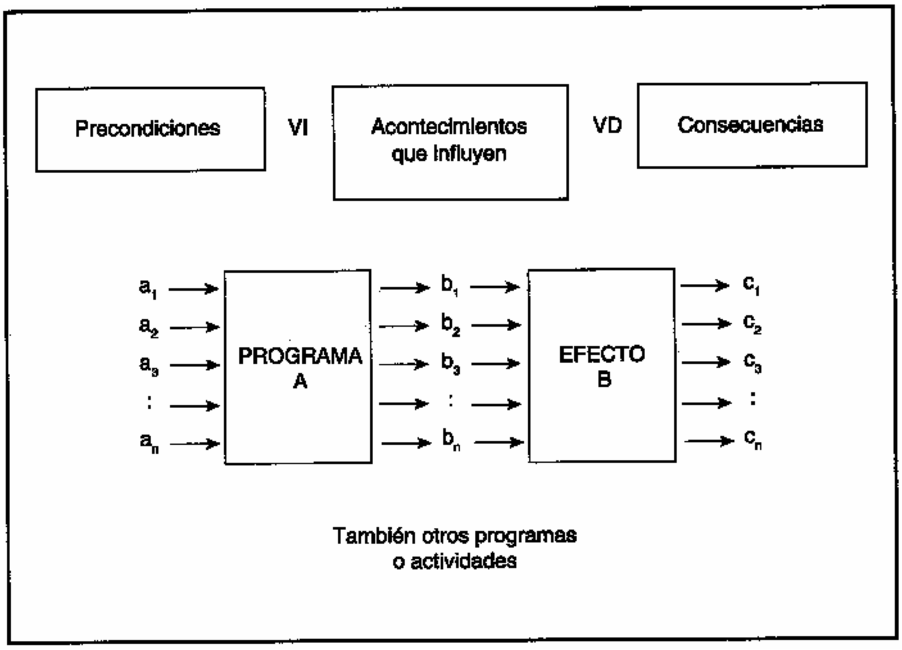


Fig. 11.4 Proceso multicasual según Suchman (1967). Adaptado de Stufflebeam y Shinkfield (1987, 121)

DISEÑOS SUSCEPTIBLES DE COMPLEMENT ARIEDAD

En muchos proyectos de investigación evaluativa es útil recoger una amplia gama de información recurriendo a datos que permitan explicar y comprender mejor el fenómeno evaluado. En este sentido es necesario conferir a la metodología la adecuada flexibilidad y adaptabilidad, con el fin de facilitar la integración de las perspectivas *empírico-analítica* y *humanístico-interpretativa.*

Algunos diseños evaluativos más susceptibles de complementariedad metodológica son los orientados a la *toma de decisiones* (Cronbach, 1963, y Stufflebeam, 1966), el diseño de *crítica artística* (Eisner, 1971) y el diseño contrapuesto (Wolf, 1974).

A) DISEÑOS MÁS ORIENTADOS A LA DECISIÓN

Los diseños de investigación evaluativa pretenden, en mayor o menor grado, ayudar a *tomar y justificar decisiones.* Sin embargo, en los diseños que ahora nos ocupan este es el cometido prioritario. Se centran en «quién» y «cómo» se deben tomar las *decisiones* y en proporcionar a los usuarios de la información el conocimiento de dónde radican los elementos determinantes de su acción educativa, qué actividades, acciones y opciones están afectadas por dichos elementos y de qué manera la información evaluativa puede incrementar el potencial educativo de los programas.

En este contexto, la investigación es utilizada tanto para el *perfeccionamiento* de un programa como para enjuiciar su valor, y el proceso de investigación suele estructurarse en tomo a las *necesidades detectadas.* Suelen plantearse cuestiones como las siguientes:

- ¿Qué es lo que hay que conseguir con el programa?
- ¿Cuál debe ser la planificación de una institución educativa determinada?
- ¿Cómo debe llevarse a cabo el plan?
- ¿Cuándo y cómo debe ser modificada la institución educativa?

Los diseños más clarificadores y representativos para el ámbito educativo son el modelo *UTOS* de Cronbach (1963) y el *CIPP* de Stufflebeam (1966).

1) Diseño UTOS

Este diseño fue elaborado por Cronbach (1963, 1980 Y 1982) Y considera la *planificación* evaluativa y las cuestiones que hay que investigar como ejes directores del diseño (Stufflebeam y Shinkfie1d, 1987):

a) *Planificación evaluativa.* El evaluador tendrá en cuenta elementos como los siguientes:

1) *Unidades.* Constituida por sujetos con determinadas características: alumnos, profesores, padres.
2) *Tratamientos.* Una unidad está expuesta a la aplicación de un tratamiento; por ejemplo, un profesor selecciona y organiza lecciones, introduce una metodología novedosa o un programa concreto.
3) *Observación.* El evaluador obtiene datos antes, durante y/o después del tratamiento. Para ello utilizará cierto tipo de tests o enviará a alguien a la clase para que observe y recoja información.
4) *«Setting».* Contexto donde se desarrolla y aplica el programa.

En una investigación concreta pueden existir varias unidades, varios tratamientos asociados y varias observaciones bajo la denominación de *uto.* En cambio, *UTO* se refiere al universo de las observaciones admisibles. La planificación ideal define un *UTO,* y luego los planes para seleccionar un *uto* que lo represente. Si la planificación es lógica y se lleva a cabo estrictamente, de los resultados de uto se pueden inferir los probables resultados de *UTO.* El evaluador piensa en *u, t yo* como en muestras de una población y controla el muestreo en la medida de lo posible para que sea representativo. Por ejemplo, el tratamiento planificado para una unidad es T, mientras que el tratamiento que realmente se aplica es t.

b) *Identificación de las cuestiones que hay que investigar.* Durante la denominada fase *divergente* se hace una lista de las posibles cuestiones, y durante la fase convergente se asignan prioridades entre ellas. Realizada una primera lista de cuestiones, deben plantearse en términos de U, T Y O que resulten significativos para los que intervendrán en la toma de decisiones. En el proceso *convergente,* algunas de estas cuestiones resultan escogidas para convertirse en *UTO* del estudio.

2) El diseño CIPP

Según Stufflebeam (1971 y 1987), el propósito más importante de la investigación evaluativa es el *perfeccionamiento* de los programas. Un objetivo fundamental del diseño *CIPP,* creado por el mismo autor, es promover el desarrollo de programas y ayudar a los directores y profesorado de una institución educativa a obtener y utilizar una información continua y sistemática. La finalidad es satisfacer las necesidades más importantes o, al menos, hacer lo posible con los recursos de que dispongan. Este diseño constituye/una estrategia adecuada que ayuda a mantener y mejorar la calidad de las operaciones de las instituciones educativas. En el modelo *CIPP* se contemplan los componentes *Contexto,* entrada o *Input, Proceso* y *Producto.*

B) DISEÑO DE CRÍTICA ARTÍSTICA

Eisner (1971 Y 1981) concibe la enseñanza como un arte y el evaluador ha de interpretar lo observado tal como ocurre en su propio contexto. Pérez Gómez *(1985b)* destaca el carácter *descriptivo* e *interpretativo* del diseño y la realización de *Juicios de valor.* A diferencia de los diseños de negociación, utiliza metáforas para describir el programa, en vez de emplear la terminología característica de las personas implicadas. Por otro lado, en la metodología cualitativa el evaluador tiende a ser inductivo, mientras que el modelo de crítica artística contempla el programa en relación a criterios de calidad previamente establecidos.

C) DISEÑO CONTRAPUESTO

A través de un proceso dialéctico entre *partidarios* y *adversarios,* este modelo trata de iluminar los aspectos más importantes del programa que hay que evaluar. Dos grupos de evaluadores investigan pros y contras de un programa para proporcionar información pertinente a quienes toman las decisiones en el ámbito educativo. Este método, utilizado por Thomas Owens, puede conjugarse con los modelos propuestos por Scriven, Stufflebeam y Provus.

DISEÑOS DE ORIENTACIÓN HUMANÍSTICO-INTERPRETATIV A

Con esta orientación los diseños utilizan una perspectiva holística y se basan en el acuerdo intersubjetivo de los participantes como criterio de evidencia. Constituyen una reacción ante las aportaciones limitadas e inadecuadas de los diseños de orientación cuantitativa para optimizar la práctica educativa, ya que los problemas que tiene que abordar el educador son demasiado complejos para ser analizados exclusivamente a través de los diseños experimentales y de las técnicas psicométricas. Los diseños, de evaluación *sin metas* (Scriven, 1967), *iluminativo* (Parlett y Hamilton, 1977), *democrático* (McDonald, 1976) y el diseño *respondente* (Stake, 1967 y 1975) utilizan las estrategias propias de la metodología dominante en el paradigma *humanístico-interpretativo* de orientación cualitativa.

A) DISEÑO SIN METAS

Scriven (1967) centra su diseño de evaluación en la valoración de los efectos y en las necesidades, prescindiendo de los objetivos del programa. Al ignorar los objetivos se evita el riesgo que impone la estrechez de unos objetivos previamente delimitados y se favorece la posibilidad de tener en cuenta efectos imprevistos, una mayor objetividad e independencia del evaluador.

La aportación innovadora de Scriven es desplazar la evaluación desde los objetivos hasta las necesidades educativas. Para él la evaluación tiene dos funciones principales: *la formativa,* que ayuda a desarrollar el programa, y la *sumativa,* que estima el valor del programa una vez que ha sido desarrollado y puesto en el mercado.

Scriven establece otra distinción entre evaluación intrínseca y final. La evaluación *intrínseca* valora las cualidades del programa, sin tener en cuenta los efectos, juzga características como las metas, la estructura, la metodología, las calificaciones y las actitudes del personal. La evaluación final no se preocupa de la naturaleza del programa o de los libros de texto, sino más bien de sus efectos. Ambos tipos de evaluación pueden contribuir tanto a la evaluación formativa como a la sumativa. Scriven considera importante la evaluación intrínseca, pero destaca que deben determinarse y juzgarse los resultados. Dados los resultados de una evaluación de necesidades, pueden juzgarse los efectos de un programa concreto como buenos, malos o indiferentes, dependiendo de si contribuyen o no a satisfacer las necesidades identificadas. Con el fin de facilitar el proceso de la evaluación, Scriven elaboró una lista de indicadores que puede consultarse en Stufflebeam y Shinkfield (1987,349).

En consecuencia, para efectuar el proceso de valoración hay que recurrir a dos tipos de información básica: los *efectos* y las *necesidades* (fig. 11.5).

Supongamos que durante la evaluación de un programa de intervención educativa hemos recogido información que permite constatar:

- el programa ejerce los efectos reales A, B, C y D;
- al planificar el programa se pretendía alcanzar los efectos C, D, E y F, pero no se habían previsto los efectos A y B, que, sin embargo, sí se han producido. En cambio, los efectos previstos E y F no se han alcanzado;
- el perfil de necesidades detectado establece las siguientes prioridades: B, C, A, D, E y F.

La medida o descripción de los efectos reales y el perfil de necesidades educativas permiten valorar la importancia de los efectos. El proceso de valoración estimará en qué medida los efectos del programa son deseables y responden a las necesidades priorizadas. El programa es útil y será valorado positivamente en relación con los efectos A, B, C y D porque cubre las cuatro necesidades más prioritarias. Algunos programas educativos pueden generar efectos no pretendidos que pudieran ser tan importantes como los objetivos iniciales. Así, en el ejemplo propuesto el efecto B no había sido previsto, pero, sin embargo, cubre una necesidad priorizada en primer lugar. Por otro lado, la valoración sería más desfavorable en relación con los efectos E y F, que cubrirían algunas necesidades si se hubieran alcanzado. Las valoraciones efectuadas deben matizarse según las implicaciones que pudieran tener en el plano institucional, con el fin de introducir los cambios adecuados que permitan optimizar la intervención.

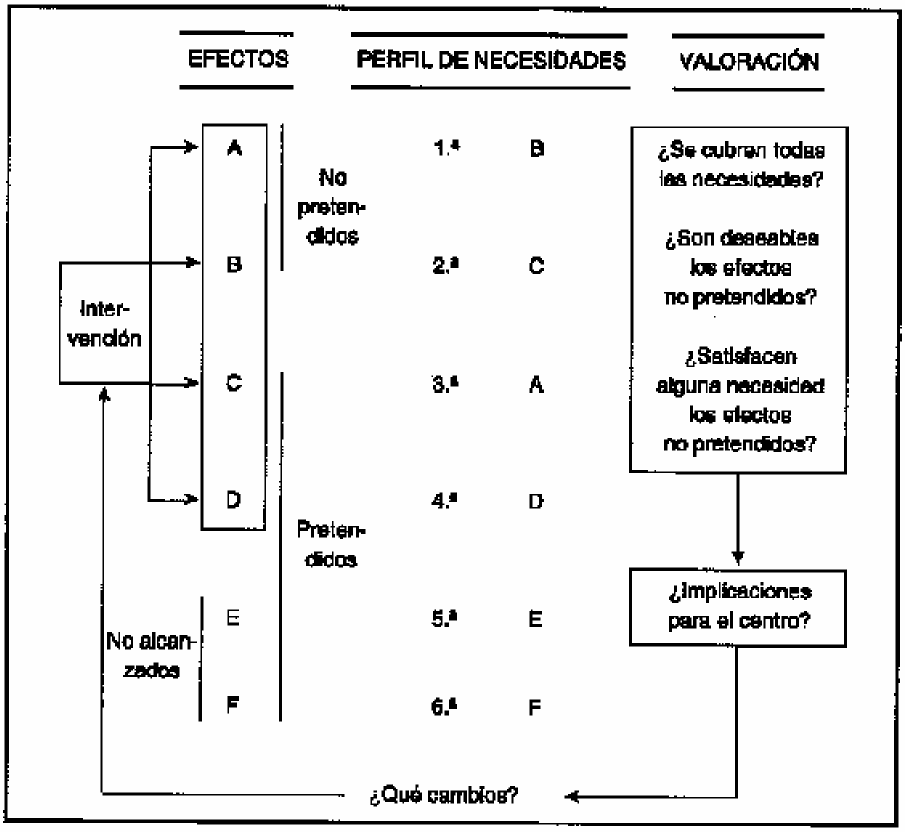


FIG. 11.5 Esquema del diseño sin metas

B) DISEÑOS DE NEGOCIACIÓN

Intentan ayudar a los educadores a que valoren y perfeccionen su contribución a la práctica educativa, orientando a las personas que realizan y aplican los programas para que efectúen sus propias evaluaciones.

Durante la aplicación de los diseños de *negociación* las distintas estrategias que hay que seguir se elaboran en función de las situaciones educativas implicadas y de común acuerdo con los participantes. Por tanto, son diseños no preestablecidos, abiertos al cambio, con el fin de no limitar las posibilidades de investigación durante el desarrollo de la evaluación. Se considera que la realidad de cada aula es singular e irrepetible; cada espacio y cada experiencia es singular, de ahí la necesidad de numerosas perspectivas y puntos de vista que el evaluador debe considerar desde una posición neutral.

Los diseños *iluminativo, democrático* y *respondente* reflejan las características mencionadas y suelen denominarse de *negociación o naturalistas.* A continuación describiremos brevemente los dos primeros y, posteriormente, abordaremos con

mayor exhaustividad el diseño respondente por ser el que refleja mejor los elementos esenciales de cualquier diseño evaluativo.

1) Diseño iluminativo

Este diseño (Parlett y Hamilton, 1977), más que centrarse sólo en la valoración del producto educativo, enfatiza el estudio intensivo del programa como *totalidad:* sus principios básicos, su evolución, sus actividades, sus logros y dificultades. Básicamente evalúan el sistema de *enseñanza* y el *medio de aprendizaje.* El *sistema de enseñanza* está constituido por planes de estudio, programas, supuestos pedagógicos, detalles acerca de las técnicas y del material. Dado que al aplicar un sistema de enseñanza (modelo abstracto) aparecerán notables modificaciones según el contexto, los alumnos y el profesor, la evaluación habrá de llevarse a cabo inmersa en las condiciones concretas donde se aplica el sistema.

El *medio de aprendizaje* se refiere al contexto material, psicológico y social en el que trabajan conjuntamente profesores y alumnos. Quedan implicadas variables culturales, sociales, institucionales y psicológicas que interactúan originando un único estado de circunstancias (presiones, opiniones, conflictos) que impregnan los procesos de enseñanza y aprendizaje.

Los problemas que hay que investigar suelen centrarse en el análisis de los procesos interactivos que tienen lugar en el aula. Su constante dinamismo origina una realidad cambiante, cuya evaluación exige la implicación y familiarización intensiva por parte del evaluador. Los problemas no pueden definirse previamente, aparecerán progresivamente durante el transcurso de los acontecimientos.

El proceso global comprendería etapas generales como las siguientes:

a) Etapa de exploración. Se observa la mayor gama posible de variables que pueden afectar al resultado del programa o de la innovación. Es una fase de reconocimiento.

b) Etapa de investigación. Se seleccionan y plantean cuestiones con el fin de elaborar una lista sistemática y selectiva de los aspectos más importantes del programa en su contexto.

c) *Etapa de explicación.* Los principios generales subyacentes a la organización del programa son especificados y se delinean posibles relaciones de causalidad.

2) Diseño democrático

Stenhouse (1984, 1987), McDonald (1976) y Elliott (1978) son considerados por Pérez Gómez *(1985b)* como los principales representantes de un modelo de evaluación que pretende desarrollar al máximo las posibilidades de los nuevos enfoques cualitativos. Estos autores ponen de relieve la necesidad de investigar y evaluar la vida del aula con planteamientos y procedimientos metodológicos naturalistas. Si la evaluación ha de servir para la realidad educativa, esta habrá de ser analizada en el marco de la realidad diaria.

La investigación trata de orientar la actividad educativa ayudando a decidir *qué* hay que hacer y cómo hacerlo. El evaluador ha de sumergirse en el contexto real y vital en que tienen lugar los acontecimientos y conocer las interpretaciones diversas que hacen los participantes que las viven directamente. De ahí que, tanto las fuentes de datos como los destinatarios de los informes, son todos los que participan en el programa. Por esta razón se denomina *diseño democrático de evaluación.* El propósito básico es facilitar y promover el cambio, llegando a una transformación real mediante la modificación de concepciones , creencias y modos de interpretar de las personas implicadas.

C) DISEÑO RESPONDENTE

Para Stake (1967, 1975 Y 1976), la evaluación es un valor observado y comparado con alguna norma y ha de enfatizar lo que se ha observado acerca del programa, detectando la satisfacción o insatisfacción que muestran las personas implicadas en el programa. La evaluación respondente contempla exhaustivamente aspectos como antecedentes, procesos o actividades del programa, sus efectos, datos para los juicios valorativos, el informe holístico y el asesoramiento a los educadores.

Teniendo en cuenta la creciente implicación de los docentes en los procesos de planificación, implantación y evaluación, el diseño *respondente*, por su carácter global, clarificador y didáctico, brinda un marco adecuado para que puedan producirse cambios reales en el ámbito educativo. En consecuencia, este diseño será objeto de una descripción más detallada.

11.5 DISEÑO RESPONDENTE

Stake (1975) elaboró un diseño centrado en responder a los problemas y cuestiones reales que se plantean las personas implicadas en un programa educativo. Según este diseño, el evaluador responde a la demanda presentada por diferentes audiencias, negociando con las personas implicadas lo que debe hacerse.

A pesar de las críticas dirigidas contra la evaluación orientada a objetivos, numerosos investigadores han seguido utilizando los objetivos como criterio de referencia, pero Stake (1967, 1975 Y 1983) introduce valiosas innovaciones. Para este autor, hay que tener en cuenta la justificación teórica o base lógica del programa que se va a elaborar, una matriz de descripción y otra matriz de juicio o valoración (fig. 11.6). El diseño exige completar estas dos matrices de información, cada una de las cuales está dividida, a su vez, en dos columnas más.

La primera columna (casillas 1, 2 Y 3) se destina a la descripción de la *planificación* del programa, y en ella se detallan los objetivos o intenciones. Esta información constituye el *modelo teórico* que se ha planificado. La segunda columna (casillas 4, 5 Y 6) se reserva para la descripción de lo que ocurre en la realidad durante la *implantación* o aplicación del programa tal como se ha planificado en la columna anterior. Esta información procedente de lo que ocurre en la realidad origina el *modelo empírico.* Las otras dos columnas abarcan los criterios o normas de calidad absolutas y relativas (casillas 7,8 Y 9) Y las valoraciones emitidas (casillas 10, 11 Y 12). Tanto

la matriz descriptiva como la de juicio contemplan antecedentes o características de la situación inicial, procesos o actividades, y efectos o resultados del programa.

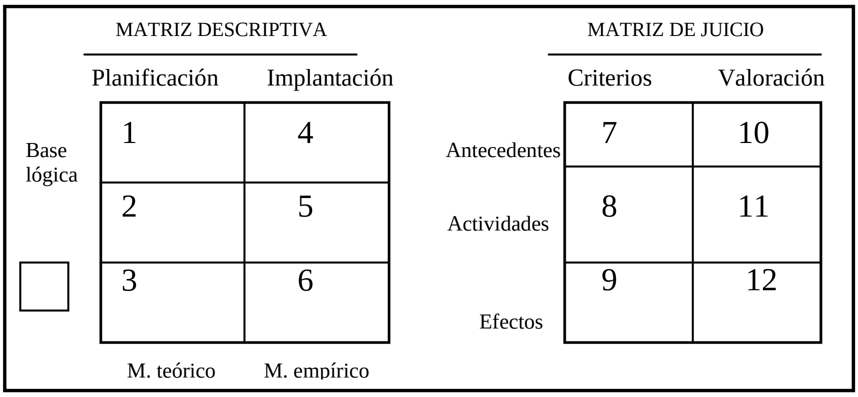


FIG. 11.6 Diseño respondente. Adaptado de Stake (1967)

Los *antecedentes* son condiciones existentes antes de que se inicie el proceso educativo o la implantación del programa y que pueden influir en los resultados. Por ejemplo, las características de las situaciones vitales que han de afrontar los sujetos que participarán en el programa. Estas características contextuales son muy cambiantes y exigirán actividades específicas, coherentes y adaptadas a dichas situaciones vitales. También puede ser necesario incluir aquí características descriptivas de los sujetos como inteligencia o rendimiento previo.

Las *actividades o procesos* didáctico s constituyen las actividades e interacciones de los alumnos con profesores, padres, tutores, compañeros y otros adultos. En general, los procesos representan el conjunto de interacciones y actividades que tienen lugar en el escenario educativo. Estas actividades planificadas constituyen el programa.

Bajo la denominación de *efectos* se incluye todo lo que se puede conseguir a través de un programa. Aquí están implicados tanto los efectos esperados como los no esperados. Los resultados son los efectos del proceso educativo y se refieren a la consecución de objetivos en relación a conductas terminales de los alumnos, a las condiciones contextuales y al comportamiento del profesor.

JUSTIFICACIÓN TEÓRICA DEL PROGRAMA

Está constituida por los planteamientos teóricos e intenciones filosóficas en que se basará el programa y sobre las que se planifica la estructura organizativa de una institución y, más en concreto, su proyecto educativo. Para enjuiciar el programa será imprescindible esta información, de ahí que tenga que ser elaborada cuidadosamente. Durante la elaboración del programa el evaluador ha de poder preguntarse: « ¿Constituye el programa que se planifica -

y se implanta un paso lógico en la realización de las intenciones básicas?».

El contenido de la *base lógica* permite analizar si la dinámica real durante la planificación e implantación del programa es coherente con las directrices y las intenciones básicas. A su vez, la *base lógica* sirve de pauta para la formulación de los objetivos del programa y también es útil para elegir grupos de referencia a los que se recurrirá para identificar los criterios de calidad y la valoración consecuente.

PLANIFICACIÓN DEL PROGRAMA

La planificación incluye los objetivos o intenciones del programa y describe todo lo que se planea, tomando conciencia de los aspectos u objetivos que se excluyen y de las razones que hay para no tenerlos en cuenta. Esta información configura el modelo teórico o conceptualización del programa y se desglosa en tres casillas temporalmente ordenadas tal como indica la numeración correspondiente a la primera columna y según los tres ámbitos ya mencionados: antecedentes o situación inicial, actividades y efectos deseados.

Casilla 1. *Previsión de la situación inicial.* Se describen los antecedentes o características previstas de los sujetos que se beneficiarán del programa y de otros que tendrán algún tipo de implicación en el mismo. También es necesario describir las características de la situación o contexto que constituirá el escenario previsto para la futura implantación de las actividades y procesos que se planifican en la casilla 2. En general, deben incluirse todas aquellas características relevantes que puedan arrojar luz sobre cuáles serán las actividades más adecuadas y que puedan condicionar los efectos esperados de la casilla 3. A modo de sugerencia puede resultar orientativo responder a preguntas como las formuladas en la figura 11.7.

Si se tuviera previsto utilizar un diseño experimental o cuasiexperimental para evaluar la efectividad de las actividades previstas, la información de la casilla 1 correspondería a la que se incluiría en el proyecto del experimento al planificar la situación pretest. En definitiva, se trata de prever cómo será la situación inicial a la que se destinará el programa. Por ejemplo, podemos planificar un programa de intervención para niños de primaria con un nivel sociocultural bajo, que estén sometidos a inmersión lingüística, obtengan puntuaciones inferiores al nivel medio de la clase en la prueba de Galí/Burt y dominen la mecánica de cálculo, pero no la comprensión del valor posicional de las cifras.

¿Qué se hace?	*¿Porqué se hace?*	*¿Quién lo hace?*	*¿Para quién?*	*¿Para qué?*	*¿Cuándo?*	*¿Dónde?*	*¿Cómo?*	*¿Con qué?*	*¿Resul tados?*

FIG. 11.7 Sugerencias para describir las características previstas para la situación inicial donde se aplicará el programa

Casilla 2. *Actividades previstas.* Dadas las características que se habrán descrito en la fase anterior, ahora se planifica cómo ha de ser el proceso educativo, qué se hará, es decir, en qué consistirá la intervención o tratamiento educativo para obtener los resultados que se incluirán en la casilla 3. Como puede apreciarse, la información de la casilla 2 correspondería a la que se incluiría en el proyecto de **un** experimento al planificar el tratamiento que se aplicará o variable independiente.

De forma orientativa, el evaluador puede formularse cuestiones como las que aparecen en la figura 11.8.

¿Qué debe hacerse?	*¿Qué puede hacerse?*	*¿Quién lo hará?*	*¿Para quién?*	*¿Para qué?*	*¿Cuándo?*	*¿Dónde?*	*¿Cómo?*	*¿Con qué?*

FIG. 11.8 Posibles cuestiones que orientan la planificación del programa

Al plantearse cuestiones como las de la figura 11.8 pueden confeccionarse esquemas como el que se reproduce en la figura 11.9 para evitar que se omitan aspectos importantes.

Por ejemplo, con el fin de mejorar el cálculo mental, el profesorado de **un** centro de primaria puede introducir juegos con dados y fichas de colores, agrupando los alumnos en grupos de cinco durante tres días a la semana. En cada casilla de la matriz de la figura 11.1 puede incluirse información relacionada con la secuenciación de las actividades.

Profesor	Grupo	Actividad	Material	Agrupación sujetos	Frecuencia duración	Objetivo

FIG. 11.9 Sugerencias para describir la planificación del programa

Así, durante la primera y segunda semanas se realizan juegos con fichas de madera coloreadas; durante la segunda semana se introducen además juegos con dados hasta la cuarta, mientras que las fichas se suprimen al acabar la segunda semana.

		Temporalización				
		1.ª semana	2.ª semana	3.ª semana	4.ª semana	etc
Actividades planificadas	juego fichas	xxxxxx	xxxxxx			
	juego datos		xxxxxx	xxxxxxxx	xxxxxxxx	
	etc.					

FIG. 11.10. Extracto de una matriz con la temporalización de actividades y el período respectivo de implantación (X)

Casilla 3. *Efectos esperados.* Por último, los resultados esperados se incluyen en la casilla 3. Esta información correspondería a la que se incluiría en el proyecto del experimento al planificar la situación postest para medir el nivel de la variable dependiente después de que se aplique la variable independiente.

Así, al introducir juegos con fichas de madera coloreadas podemos esperar que los alumnos comprendan mejor el valor posicional de las cifras, lo que favorece que los alumnos asimilen recursos que les permitan agilizar el cálculo mental a través de juegos con dados. En la figura 11.11 se registran horizontalmente los tipos o niveles de efectos esperados y verticalmente las actividades que se han planificado con el fin de alcanzar dichos resultados. Sin embargo, generalmente, no se espera que las distintas actividades contribuyan por igual al logro de cada efecto previsto, de ahí la conveniencia de incluir símbolos (+ -) en la matriz para indicar una estimación de la contribución relativa de cada actividad a los distintos efectos. También es muy sugerente anotar en las casillas de la misma matriz o de otra preparada al efecto los comentarios de las personas implicadas o de expertos en cada ámbito, en torno a las expectativas que mantienen acerca de la posible contribución de cada actividad para lograr cada efecto propuesto.

		Tipos o niveles de efectos esperados				
		Compren-sión	Agilidad cálculo	Motivación	Valor posicional	etc.
Actividades planificadas	juego fichas	++	+	+++	+	
	juego dados	–	+	+	–	
	3.ª	Comentarios sobre la relación de causalidad prevista entre las actividades planificadas y los efectos esperados				
	4.ª					
	etc.					

FIG. 11.11. Extracto de una matriz con actividades y efectos previstos al planificar un programa educativo para optimizar el cálculo mental

IMPLANTACIÓN DEL PROGRAMA

Para elaborar la planificación puede ser necesario recurrir ocasionalmente a ciertas técnicas de recogida de datos, ya sea con el fin de describir mejor la situación a la que se tiene previsto aplicar el programa, o bien para recoger opiniones sobre la conveniencia de determinadas actividades y sus posibles efectos.

Sin embargo, para describir lo que ocurre durante la implantación del programa, al describir aspectos como la situación real de partida, cómo se realizan las actividades y los efectos que se van detectando, es imprescindible contar con una información tan exhaustiva y variada que es preciso recurrir a un proceso interactivo en la utilización de las técnicas de recogida de datos. La compleja relación que existe entre técnicas como la observación, la entrevista, el análisis de documentos y la aplicación de instrumentos estructurados puede esquematizarse según la analogía del *missile*

(Wemer y Schoepfle, 1987). El investigador completa la información recogida con unas técnicas acudiendo a los datos aportados por otras técnicas con el fin de aproximarse progresivamente a la realidad educativa objeto de evaluación (fig. 11.12).

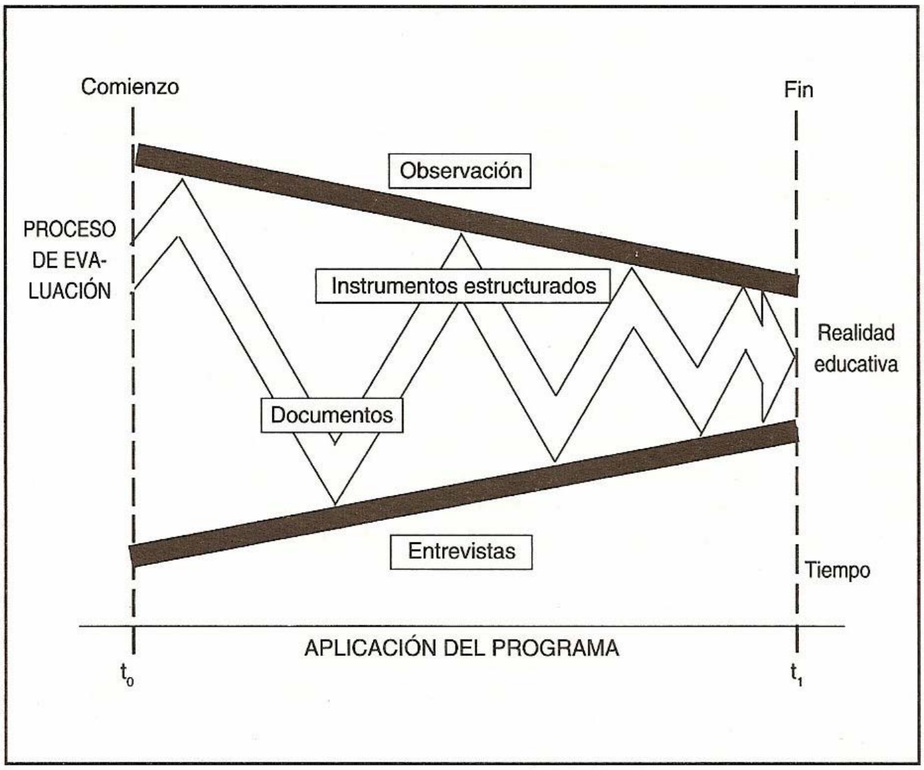


FIG. 11.12 Proceso interactivo para evaluar la realidad educativa.
Adaptado de Wemer y Schoepfle (1987)

Casilla 4. *Características reales de la situación inicial.* Se describen los antecedentes o características de la situación inicial real donde se implanta el programa, de los sujetos que se beneficiarán directamente del mismo y de otros que tengan algún tipo de implicación. En general, deben incluirse características que nos permitan estimar en qué medida el contexto real de la situación inicial coincide con el previsto, ya que las actividades que se aplicarán son coherentes con la situación inicial prevista.

Si se establece un paralelismo con el diseño experimental, esta información se corresponde con la que aporta la aplicación del pretest. En definitiva, se trata de constatar que la situación inicial a la que se destina el programa coincide con la que se había previsto. Es decir, siguiendo con el ejemplo que habíamos propuesto, los alumnos habrán de ser de primaria y procederán de un nivel sociocultural bajo, estarán sometidos a inmersión lingüística, se constatará que obtienen puntuaciones inferiores

al nivel medio de la clase en la prueba de Galí/Burt y que dominan la mecánica de cálculo, pero no la comprensión del valor posicional de las cifras.

Casilla 5. *Aplicación de las actividades previstas.* Ahora se describe cómo es el proceso educativo, qué se hace, es decir, en qué consiste la intervención o tratamiento educativo. En lo que se refiere al experimento, equivale a la fase de aplicación de los tratamientos o variables independientes.

Las preguntas sugeridas en la casilla 2 también pueden ser de utilidad ahora. El evaluador se puede preguntar: ¿qué se hace durante la aplicación?, ¿quién lo hace?, ¿a quién va dirigido?, ¿para qué lo hace?, ¿cuándo, dónde, cómo y con qué materiales se realizan las actividades? Siguiendo con el ejemplo del cálculo mental y en la medida de lo posible, hay que describir la actuación de los profesores y de los alumnos durante el juego con las fichas y con el dado, en qué consisten las actividades que se realizan con estos materiales, registrando características como el color y las dimensiones, la agrupación de alumnos, la frecuencia o duración de las sesiones y el objetivo que pretende cubrir el profesor. También debe incluirse información sobre la temporalización de las actividades.

Casilla 6. *Efectos reales que se producen.* Recordemos que, teniendo en cuenta la necesidad de utilizar distintas fuentes de datos, ya se han descrito los antecedentes reales o condiciones ya existentes antes de iniciar las actividades del programa, los procesos, actividades o interacciones que se dan en la realidad educativa entre los participantes implicados. Ahora se describirán los efectos o resultados que se van produciendo durante la progresiva aplicación del programa, generalmente en ámbitos como capacidades, logros, actitudes y aspiraciones. El concepto de efectos se toma aquí en sentido amplio y abarca el impacto que causa en las personas implicadas, los resultados evidentes y los confusos, los buscados y los no buscados, los de largo y los de corto alcance.

	EFECTOS A CORTO PLAZO (+ −)	EFECTOS A LARGO PLAZO (+ −)	EFECTOS NO PRETENDIDOS (+ −)
Visto por alumnos			
Efectos en profesores	Aumenta interés • • •	• • •	Mejora imagen centro • • •
Efectos en alumnos	• • *Conflictos	Creatividad • •	Dedicación • •
Visto por profesores			
Efectos en profesores	• • •	• • •	• • •
Efectos en alumnos	Mucha motivación • • •	• • •	Indisciplina • • •

—— = Opinión expresada con énfasis por una persona o por más de una persona
* = Inferencia hecha por el evaluador

FIG. 11.13 Matriz de efectos desde la perspectiva de profesores y alumnos

Para constatar los resultados o efectos de un proceso de intervención puede recurrirse a una matriz como la que hemos expuesto al describir los efectos esperados (fig. 11.11). También puede utilizarse una matriz similar a la que aparece en la figura 11.13. En ella se incluyen dos fuentes de información (profesores y alumnos) y efectos a corto y a largo plazo, así como efectos no pretendidos.

ANÁLISIS DE LA INFORMACIÓN

Recordemos que en el proceso de evaluación de un programa hemos considerado, según el diseño de Stake, la *base lógica* y las descripciones de lo *planificado* y de la *implantación. La planificación constituía* la *conceptualización* del programa o *modelo teórico,* mientras que la descripción de la *implantación* daba lugar al *modelo empírico.* Ahora debe realizarse el análisis de la coherencia entre la información correspondiente a las casillas 1,2,3,4,5 Y 6 con el fin de enjuiciar posibles *discrepancias* a la luz de los *criterios* valorativos como base para efectuar una *valoración* progresiva del programa. Aunque Stake (1967) utiliza la denominación de *congruencias y contingencias,* en lo sucesivo utilizaremos el concepto genérico de coherencia con el fin de facilitar la comprensión del análisis.

A) SECUENCIA DEL PROCESO DE ANÁLISIS

El proceso puede iniciarse estimando la *coherencia* entre los elementos que constituyen la *planificación.* Existe coherencia cuando los efectos pretendidos pueden considerarse como una consecuencia lógica de los antecedentes y los procesos o actividades que se han planificado.

Admitida la *coherencia en la planificación,* es decir, la correspondencia lógica y la plausibilidad del planteamiento teórico o conceptual del programa, el análisis se centra en el grado de *coherencia entre la planificación y la implantación del programa.*

Aceptada la coherencia en el modelo teórico y constatada también la coherencia entre los modelos teórico y empírico, debe aceptarse necesariamente la *coherencia* durante la *implantación* o coherencia empírica. Dados los antecedentes descritos en la realidad, los efectos observados se atribuyen a las actividades y procesos implantados. Sin embargo, cuando el modelo teórico no ofrece las debidas garantías de validez y/o existen incoherencias o discrepancias entre los modelos teórico y empírico, está más indicado completar el análisis estimando también la coherencia empírica. Para una mayor clarificación desglosaremos la secuencia del proceso.

B) COHERENCIA EN LA PLANIFICACIÓN

Con el análisis de la *coherencia en la planificación* se identifican, a nivel conceptual o teórico, los efectos asociados a antecedentes concretos y a las actividades. Para ello se lleva a cabo un análisis lógico con el fin de analizar y valorar las implicaciones mutuas entre antecedentes propuestos y actividades planeadas, y entre estas actividades y los efectos esperados. Las tres casillas que forman la columna de

lo planificado constituyen una fase teórica de carácter deductivo, donde los resultados, lógica y necesariamente, son efectos de la interacción entre las condiciones antecedentes y las actividades que hay que implementar (fig. 11.14).

Para determinar si la planificación es coherente hay que recurrir a marcos teóricos, a la revisión bibliográfica sobre el tema, a la consulta de expertos y al consenso entre los participantes implicados. Además del conocimiento teórico, se recurre a la experiencia previa y a intuiciones que puedan aportar informes evaluativos de programas similares. Así, en el ejemplo que hemos comentado de un programa para optimizar el cálculo mental nos preguntaremos: ¿serán adecuados los juegos con dados y fichas coloreadas de madera para alumnos de primaria de nivel sociocultural bajo? Según teorías psicológicas como la cognitivo-evolutiva y las teorías sobre el juego, ¿puede esperarse que aumente la comprensión del valor posicional de las cifras con las actividades tal como se han planeado? ¿Aumentará la agilidad para el cálculo?

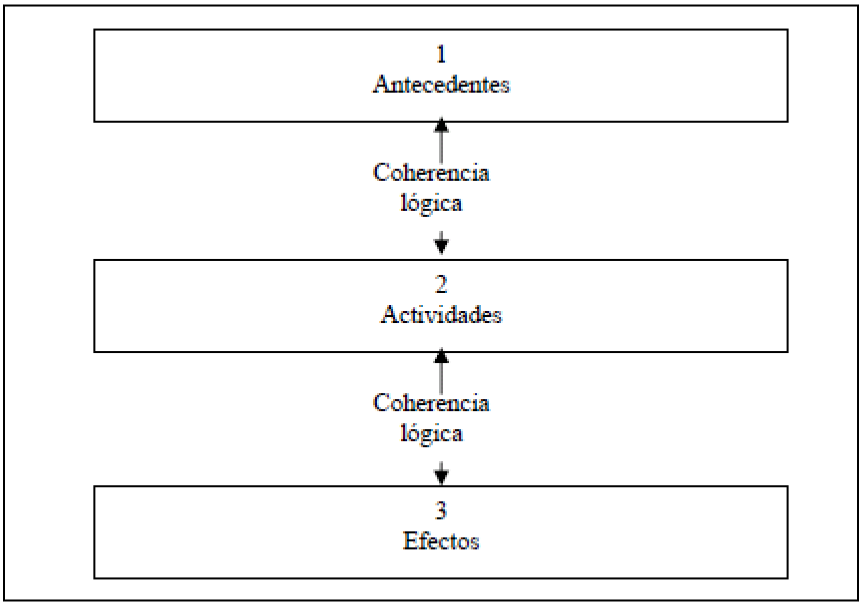


FIG. 11.14 Proceso para estimar la coherencia en la planificación

C) COHERENCIA ENTRE PLANIFICACIÓN E IMPLANTACIÓN

Este análisis trata de estimar hasta qué punto el proceso de implantación responde a la planificación. Es decir, detecta si los objetivos o intenciones planeados en el modelo teórico se cumplen en la realidad tal como se habían previsto, analizando discrepancias entre lo planificado y lo que realmente sucede durante la implantación (fig. 11.15). La coherencia debe existir entre antecedentes previstos y observados, entre actividades planeadas y realizadas, y entre efectos esperados y observados.

Para cada objetivo especificado en las casillas 1, 2 Y 3 del modelo teórico o conceptual del diseño se recoge información, como pueden ser las percepciones y vivencias de los participantes, para estimar el grado en que se observaba o se alcanza en la realidad lo que se había planificado. Así, se recaba información para ver si las condiciones antecedentes se cumplen tal como se especifican y si las actividades o procesos se desarrollan con arreglo a lo previsto. De la misma manera se abordan los posibles efectos.

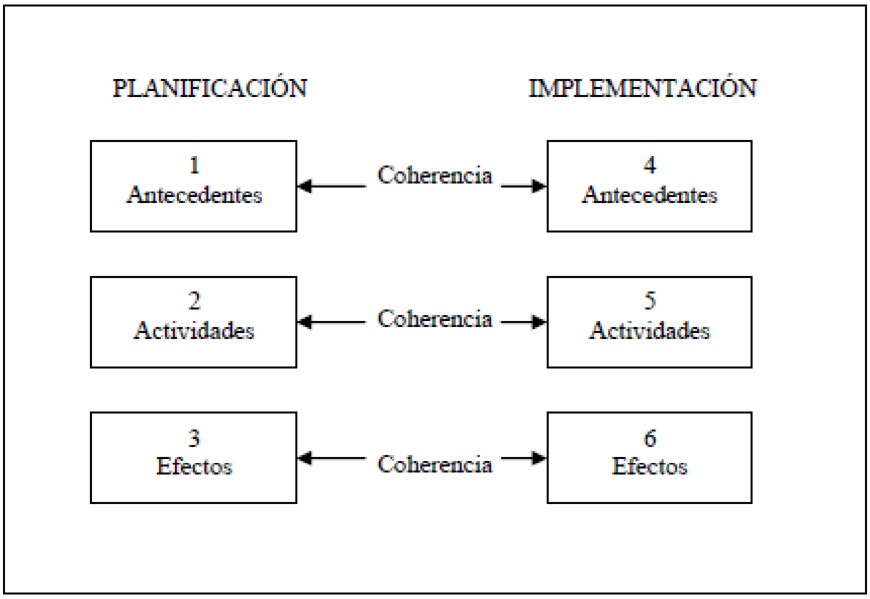


FIG. 11.15. Proceso para estimar la coherencia entre planificación e implantación

Como puede apreciarse, este análisis determina el grado en que el modelo empírico o aplicado se ajusta al modelo teórico pretendido, y por ello descubre posibles deficiencias de aplicación.

D) COHERENCIA DURANTE LA IMPLANTACIÓN

El análisis de la coherencia durante la implantación (fig. 11.16) ha de basarse, en la medida de lo posible, en criterios de evidencia empírica. Las relaciones de causalidad entre antecedentes, actividades y efectos pueden buscarse desde la perspectiva cuantitativa, con diseños experimentales, cuasiexperimentales o incluso ex-post-facto, pero también desde un punto de vista cualitativo, a partir de las percepciones de los propios participantes.

Cuando las actividades y los efectos no están especificados, deberán identificarse a partir de los datos recogidos (fig. 11.17). La clasificación cruzada de cualquier actividad con alguno de los efectos del programa origina una celda en la matriz. Así, la primera casilla ha sido originada por la intersección entre la actividad 1 y el efecto a.

Implantación del programa
Modelo empírico o aplicado

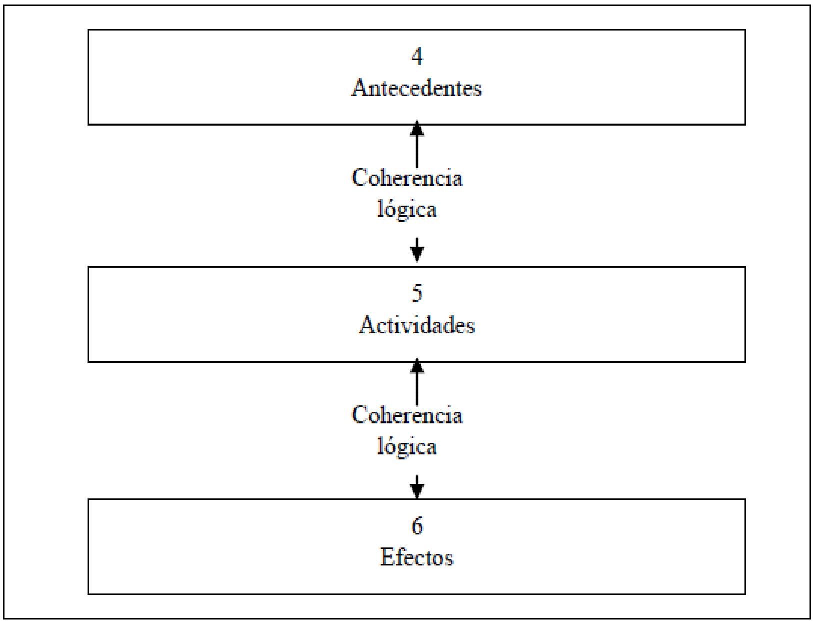


Fig. 11.16 Proceso para estimar la coherencia durante la implementación

En cada casilla se incluye todo lo que ayude a entender la relación entre las actividades y los efectos respectivos.

		Tipos o niveles de efectos esperados				
		Compren-sión	Agilidad cálculo	Motivación	Valor posicional	etc.
Actividades planificadas	juego fichas	++	+	+++	+	
	juego dados	–	+	+	–	
	3.ª	Comentarios sobre la relación de causalidad prevista entre las actividades planificadas y los efectos esperados				
	4.ª					
	etc.					

Fig. 11.17 Extracto de una matriz con actividades y efectos de un programa educativo

La información recogida con esta matriz proceso-producto debe compararse con su homóloga de procesos y efectos planificados, tal como ha sido expuesta al describir los efectos esperados que se incluyen en la casilla 3.

Con el fin de clarificar mejor las contingencias o relaciones de causalidad, en el modelo teórico, y especialmente en el empírico, se pueden elaborar y completar matrices como las que se expondrán en la parte IV sobre el análisis de datos cualitativos. Constatar la relación de causalidad entre actividades y efectos son aspectos fundamentales en la elaboración de programas. En las instituciones puede faltar una especificación de los efectos, que puede ser debida a una necesidad de individualizar o personalizar el proceso. Otras veces la naturaleza del programa hace que los efectos sean totalmente imprevisibles y, en muchos casos, obedece a que las actividades no han sido planeadas cuidadosamente. En estos casos es especialmente útil elaborar las matrices proceso-producto (Miles y Huberman, 1987, y Patton, 1987) para organizar las actividades del programa, sus efectos y su relación o falta de relación.

Cuando no se detecta la presencia de un efecto esperado se proponen hipótesis en torno a posibles causas que motivan la discrepancia. Para plantear y contrastar las hipótesis el evaluador retrocede en la matriz descriptiva y recurre a datos sobre antecedentes y sobre actividades que puedan arrojar luz sobre dichas hipótesis. El objetivo es determinar si la ausencia de un efecto pretendido es debida a una falta de contingencia lógica o de congruencia, es decir, a deficiencias en el modelo teórico o en el aplicado. Como es evidente, la distinción tiene repercusiones importantes para la toma de decisiones, ya que las deficiencias teóricas exigen una reconceptualización, mientras que las deficiencias de aplicación exigen una optimización en la implementación de lo planificado.

E) CONSIDERACIONES EN TORNO A LA COHERENCIA

Si se demuestra la coherencia lógica en el modelo teórico (planificación) y existe coherencia entre los modelos teórico (planificación) y empírico (implantación), debe

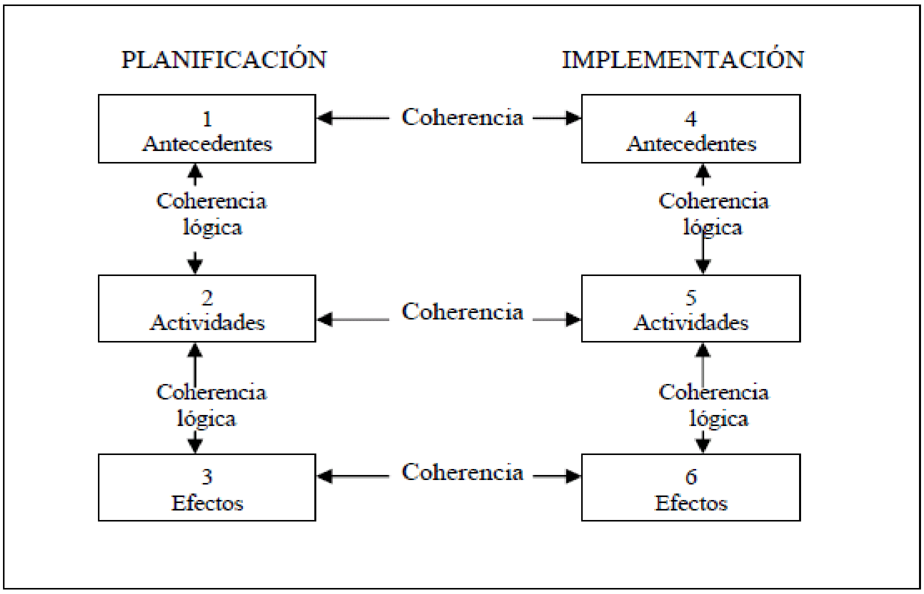


Fig. 11.18 Análisis de la información. Adaptado de Stake (1967)

existir también coherencia en el modelo empírico, es decir, durante el proceso de implantación. En ese caso, los efectos observados se pueden atribuir a las actividades implementadas en una situación educativa con determinadas características o antecedentes. La coherencia teórica y la coherencia entre los modelos teórico y empírico constituyen las condiciones necesarias y suficientes para validar la inferencia evaluativa de que los efectos esperados están siendo alcanzados satisfactoriamente y que pueden atribuirse a la incidencia del programa.

Ahora bien, en la práctica, es difícil tener plena seguridad de la existencia de coherencia lógica y de coherencia entre los modelos teórico y empírico. De ahí la necesidad de completar el análisis estimando también la coherencia empírica durante la implantación del programa. Como consecuencia, en la medida de lo posible, puede convenir un análisis global como el de la figura 11.18.

Al contemplar los tres tipos de coherencia que aparecen en la figura 11.18 es posible delimitar mejor verdaderas relaciones de causalidad (Shapiro, 1987) entre antecedentes observados, actividades realizadas y efectos alcanzados.

En definitiva, la existencia conjunta de coherencia teórica y de coherencia entre los modelos teórico y empírico implica coherencia empírica, es decir, la inferencia evaluativa de que los efectos reales observados son debidos a la interacción de los antecedentes observados y las actividades observadas.

CRITERIOS DE CALIDAD

Aunque los distintos componentes del diseño no siguen un proceso secuencial, pues en la realidad se complementan unos a otros e interactúan entre sí, de modo indicativo hemos descrito la planificación, la implantación del programa y la coherencia

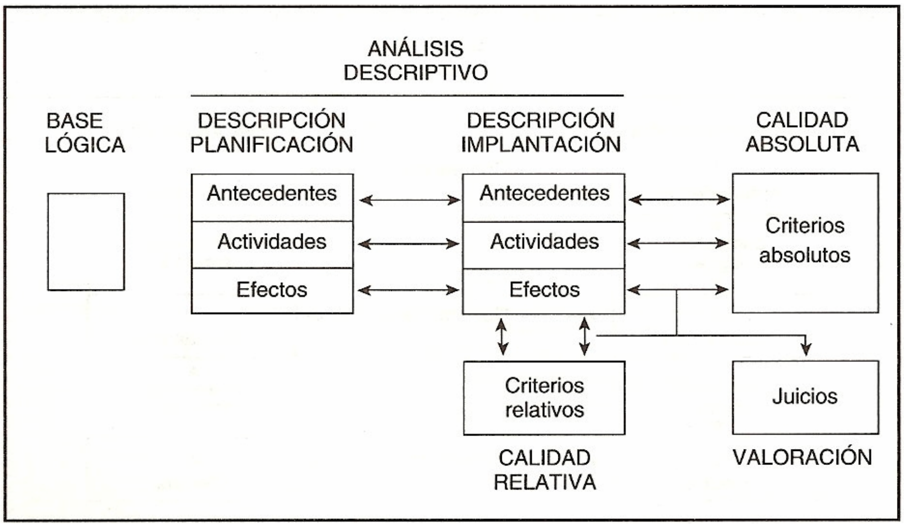


FIG. 11.19 Aspectos que hay que considerar para evaluar un programa. Adaptado de Stake (1976)

entre sus elementos. Recordemos ahora que la matriz de juicio estaba constituida por criterios de eficiencia educativa (casillas 7, 8 Y 9) que permiten efectuar las valoraciones que completarán la matriz de juicio (casillas 10, **11** Y 12). Estos criterios son cotas, puntos de ejecución o niveles de desarrollo que tienen un valor referencial ampliamente reconocido y constituyen criterios explícitos para valorar la calidad de la oferta educativa del programa. Hay que recurrir a criterios absolutos y relativos. Así, la matriz de juicio de la figura 11.6 ha sido desglosada originando la figura 11.19.

A) CALIDAD ABSOLUTA

La calidad absoluta se estima a partir de los criterios absolutos. Estos representan un conjunto de especificaciones ideales tal como son establecidas por grupos de referencia adecuados, por ejemplo, filósofos, psicólogos, pedagogos y profesores.

Los criterios absolutos engloban las convicciones personales acerca de los efectos deseables para un programa de intervención educativa. Como puede haber muchos grupos referenciales, también puede haber diferentes grupos de criterios, originando las denominadas normas o estándares de calidad. Aquí, el modelo de Stake asume el supuesto de la convergencia de los valores educativos, tiende a ignorar el pluralismo y sostiene implícitamente que el consenso es viable. Así, en el contexto de nuestro ejemplo sobre optimización del cálculo mental y según Fernández Quiles (1988), autores como Constance Kamii, investigadora de las teorías de Piaget, plantean que es absurdo exigir a los alumnos de primero de EGB que aprendan el valor posicional de las cifras en el número, ya que dicho aprendizaje supone un nivel de maduración mental superior. Sin embargo, en las programaciones de muchos países se exige este contenido en estas edades.

B) CALIDAD RELATIVA

La calidad relativa se estima recurriendo a criterios relativos que proceden de programas con objetivos similares al que se evalúa. La calidad relativa exige comparar con otros currículos alternativos, cuyos programas se admiten como satisfactorios. Es imprescindible recoger datos descriptivos de dichos programas y extraer las normas

TABLA 11.2 Extracto de cuestiones formuladas a los alumnos de cinco C. P. de Comella de Llobregat (Barcelona) durante un seminario organizado por el FOPI

Cuestión 1 Coloca los números 46 y 7 uno sobre otro y súmalos
Cuestión 2 Haz lo mismo con los números 23, 51 Y 2
Cuestión 3 Escribe las decenas y unidades que tiene el número 35
Cuestión 4 Escribe que número forman 4 decenas, 1 centena y 7 unidades
Cuestión 5 Escribe el número más grande que puedas con las cifras 3, 1 Y 6 sin repetirlas

relativas para compararlas con el programa que hay que evaluar. A modo de ejemplo podemos citar los resultados obtenidos en cinco colegios públicos de Comella de Llobregat (Barcelona) implicados en el Seminario de Ciencias y Matemáticas del FOPI cuando se pidió a los alumnos que respondieran a cuestiones similares a las de la tabla 11.2 (Femández Quiles, 1988).

Después de tabular las respuestas de los alumnos se obtuvo la tabla 11.3, que puede servir de criterio relativo para centros de características similares a los que participaron en el FOPI.

TABLA 11.3 Porcentaje de alumnos que resuelven bien cada cuestión formulada en la entrevista en un colegio de Comella de Llobregat (Femández Quiles, 1988)

	Cuestión 1	*Cuestión 2*	*Cuestión 3*	*Cuestión 4*	*Cuestión 5*
2.º EGB	83	84	76	40	67
3.º EGB	76	85	75	59	78
4.º EGB	94	94	81	64	87
5.º EGB	93	93	90	77	99

PROCESO DE VALORACIÓN

La fase de valoración parte de la base lógica, y contempla posibles discrepancias o incoherencias en la planificación, entre planificación e implantación y durante la implantación a la luz de una diversidad de criterios de calidad o deseabilidad educativa. Las discrepancias surgidas en cualquiera de los tres ni veles (antecedentes, actividades y efectos) han de confrontarse con los criterios o normas pertenecientes a la primera columna de la matriz de juicio.

Juzgar o valorar es asignar un peso a cada grupo de criterios con el fin de obtener una estimación del mérito del programa a partir de las comparaciones relativas y absolutas que se hayan realizado. El proceso de valoración vendrá matizado por los criterios absolutos y relativos. Así, la valoración de un programa para optimizar el cálculo mental basado en la comprensión del valor posicional de las cifras, dependerá de la opción que tomemos en el conflicto planteado más arriba, en tomo a la capacidad mental que asumamos en los alumnos según la teoría piagetiana o según la que subyace en las directrices de programas de algunos países. De la misma manera, se tendrá también en cuenta la información procedente de criterios relativos como puede ser el porcentaje de alumnos que resuelven cada ejercicio señalado en el apartado anterior.

IMPLICACIONES PARA LA TOMA DE DECISIONES

Con el diseño respondente de Stake pueden detectarse efectos indeseados debidos a una *falta de coherencia lógica* en el modelo teórico o conceptual que subyace en el

programa y/o debidos a que la *aplicación real del programa no es coherente* con el modelo conceptual previo. Para clarificar mejor la torna de decisiones conviene diferenciar las implicaciones que se desprenden de las deficiencias en los niveles *conceptual* y *aplicado* (Shapiro, 1987).

Un programa educativo basado en un *planteamiento teórico deficiente* indica que con los procesos o actividades implicadas no se podrán alcanzar los resultados deseados, pues su consecución viene mediatizada por falsas relaciones hipotetizadas entre antecedentes y procesos y entre procesos y efectos. La torna de decisiones podrá articularse en tomo a dos opciones:

1) Elaborar un nuevo programa o modificar el existente para alcanzar los objetivos propuestos.
2) Modificar los objetivos para que sean coherentes con los resultados obtenidos.

En cambio, las *deficiencias de aplicación* exigen que el programa se operativice de forma más coherente con el modelo conceptual propuesto. Sólo cuando se consigue esta coherencia será viable contrastar el modelo con los datos empíricos. Corno es obvio, las deficiencias debidas a la falta de coherencia entre el programa planificado y el que se aplica en la realidad no invalidan necesariamente el modelo conceptual subyacente.

11.6 POSIBILIDADES Y LÍMITES

Los diseños de orientación *empírico-analítica* son más viables para evaluar efectos finales y cuantificables, siendo menos indicados cuando el objetivo prioritario es perfeccionar el programa. En ocasiones se establece un control estricto de las variables intervinientes, pero a costa de la artificialización del contexto.

Los diseños que se han expuesto en el marco de la perspectiva humanístico-*interpretativa,* por ser más flexibles y sensibles ante acontecimientos imprevistos y el cambio, son los que mejor responden a las exigencias de la investigación acción, en el sentido de que ayudan a los educadores a elaborar, implantar y evaluar sus propios programas, sin restringir el análisis a lo observable y cuantificable o al producto final. Sin embargo, sus principales deficiencias radican en la posibilidad de que alguna parte interesada manipule la evaluación en detrimento de la autenticidad, en la falta de credibilidad externa y en la confidencialidad.

Los diseños más susceptibles de *complementariedad* metodológica proporcionan una base lógica que ayuda a los profesionales a responder de las decisiones tornadas durante la aplicación de un programa. Sin embargo, estos estudios suelen ir dirigidos a los niveles administrativos más que alas educadores de los centros, ya que existe más autoridad decisoria en los niveles más altos.

El mayor problema con la clasificación de los diseños que se ha propuesto radica en su interpretación y aplicación. Las tipologías de diseños, que en conjunto son resultado de variar el énfasis en los propósitos, métodos y componentes del proceso evaluativo, en la práctica suelen convertirse en fuente de confusión. En consecuencia, didácticamente, es imprescindible proponer un diseño clarificador y lo más global posible. El diseño respondente de Stake (1967 y 1976) puede reunir estos requisitos.

Por otro lado, este diseño puede aglutinar las estrategias de los demás diseños enmarcados en la perspectiva humanístico-interpretativa, e incluso de la orientación empírico-analítica.

En general, el diseño *respondente* presenta la ventaja (Guba y Lincoln, 1982) de ampliar el concepto de objetivos incluyendo factores contextuales, habiendo sido aplicado en el ámbito de la medicina (Shapiro, 1987) y en el educativo (Del Rincón y Santolaria, 1989). Este diseño ofrece un retrato completo y holístico de las actividades del programa enfatizando la descripción de los procesos, destacando el juicio como aspecto más fundamental de la evaluación y ofreciendo información para derivar criterios valorativos absolutos y relativos. Por otro lado, es suficientemente flexible para posibilitar planteamientos eclécticos. Sin embargo, el diseño *respondente* no explica detalladamente la manera de elaborar los criterios, ofreciendo sólo mínimas directrices para operativizarlos, mientras que, en la práctica, surgen dudas sobre el tratamiento de los valores en conflicto.

CAPÍTULO 12

Investigación-acción

12.1 PERSPECTIVA HISTÓRICA

a) ORÍGENES DE LA INVESTIGACIÓN-ACCIÓN

La mayor parte de autores de la historia de la educación coinciden en atribuir a la obra de Dewey (1929) *Sources of the Science of Education* la paternidad de las fuentes próximas de la investigación-acción. Sus ideas sobre la «pedagogía progresiva»: el Cirraéter democrático de la educación, el aprendizaje en la acción, la necesidad de implicación de los maestros en los proyectos de investigación; en definitiva, su pensamiento crítico y reflexivo y sus ideas de democracia y participación subyacen implícita o explícitamente en los planteamientos de la investigación-acción (Schubert, y Schubert, 1984). La investigación-acción es un resultado directo y lógico de a posición progresiva en educación; después de enseñar a los niños cómo trabajar juntos para resolver sus problemas, el siguiente paso era que los maestros adoptaran los métodos que habían enseñado a los niños y aprendieran a resolver sus problemas de manera cooperativa (Hodgkinson, 1957).

A las ideas de Dewey se unen como fuentes de la investigación-acción los escritos de Collier (1933-1945), comisario de la oficina de asuntos indios, quien utiliza la expresión *action research* y señala la necesidad de aplicarla a los programas sociales y la colaboración entre investigador práctico y cliente.

Si bien las raíces próximas de la investigación-acción se adscriben a las ideas de Dewey y Collier, es justo reconocer, y así lo consideran la mayoría de autores, que la aportación de Lewin (1946) a la investigación-acción fue decisiva (Thirion, 1980). Este psicólogo social, que acuñó la expresión *Action Research* en el año 1946, la emplea para describir un enfoque de investigación, que sin romper con el modelo tradicional empírico-analítico de la psicología, supone una adaptación de dicho modelo a los programas de acción social (Pini, 1981). La originalidad de Lewin radica en la aportación de un nuevo concepto de investigación. Para este autor, en palabras de su bibliógrafo Marrow (1969), la investigación es ante todo investigación-acción. «No queremos acción sin investigación, ni investigación sin acción.» Mediante . la investigación-acción, señala Lewin, los avances teóricos y los cambios sociales se pueden lograr simultáneamente.

La investigación-acción fue absorbida por la educación casi tan pronto como surgió. El mismo Lewin colaboró en algunos proyectos de desarrollo curricular. Sin embargo, sería Carey (1949), decano del Teachers College de la Universidad de Columbia,-quien llevó las ideas de Lewin al campo de la educación y se convirtió en su principal defensor. A través del Instituto Horace Mann-Lincoln realizó varios proyectos de investigación, generó gran número de informes y publicó obras tan significativas como *Action Research to Improve School Practices* y *Curriculum Development through Action Research.* Junto a la obra de Carey es justo señalar los trabajos de Taba y Noel (1957) y de Shumsky (1958) como pioneros en el área de investigación-acción educativa.

b) DECLIVE DE LA INVESTIGACIÓN-ACCIÓN

Si el período entre 1944 y 1953 supone el mayor impacto de la investigación-acción en el área de la educación, y como señala G. Joncich (1973), más de la mitad de los artículos que tratan de metodología de las ciencias de la educación son clasificados bajo la rúbrica de investigación-acción, el período entre 1953 y 1957 representa la pérdida de. interés por la investigación-acción, al menos en el área educativa (Kernmis, 1983).

En la década de los cincuenta aparecen las primeras críticas contra la investigación-acción. Wiles (1953) pide una mayor precisión en la definición del concepto de investigación-acción, y Hodgkison (1957) critica su' metodología, el tiempo que consume, su carácter «acientífico» según los cánones de la filosofía de la ciencia social, a la vez que discute la posibilidad de que los profesores puedan investigar,

Durante los años sesenta la investigación-acción tiene poca incidencia en la educación, tal vez debido a la emergencia del diseño «R y D» (Investigación y Desarrollo) que se impone con fuerza en Estados Unidos. En opinión de Escudero (1987), como consecuencia de la importancia que adquirieron los modelos de tipo positivista y los esquemas preferentemente tecnológicos para el diseño y desarrollo de las innovaciones educativas, así como para la formación de los profesores.

c) RESURGIMIENTO DE LA INVESTIGACIÓN-ACCIÓN

En la década de los setenta resurge el movimiento de investigación-acción en el contexto de colaboración entre maestros e investigadores en el desarrollo del currículum; primeramente dirigido al problema de cómo realizar los valores fundamentales educativos en la acción. Es impulsado por investigadores vinculados a proyectos de investigación, tales como Stenhouse (1970), Elliott (1973), Allal, y otros (1979), que presentan modelos alternativos a la investigación tradicional educativa. Estos proyectos utilizaban los principios de la investigación-acción. Como plataforma del movimiento de investigación-acción se constituye en el Instituto Cambridge el órgano de expresión CARN (Classroom Action Research Network), que bianualmente publica un tema monográfico referente a la investigación-acción. Este movimiento se ramifica en varias corrientes, unas puramente geográficas y otras de tipo más conceptual. Diferentes escuelas radicadas en Francia, Suiza, Alemania y en Estados Unidos, muestran

la fuerza del movimiento. Adquiere una considerable pujanza en países como Australia, de forma particular el Educational Action Research Group fundado por Kemmis en la Universidad de Deakin, y el grupo de Montreal, con el profesor A. Morin.

En lo que se refiere a España, De la Orden (1967) publica un artículo, «La investigación educativa y su organización dentro del sistema escolar español», donde utiliza la expresión *investigación activa* por *Action Research* y en otras ocasiones *investigación operativa;* estas denominaciones se han ido utilizando con diferentes matices tanto en España como en Francia, acabando en la actualidad imponiéndose la expresión *investigación-acción.*

Aunque antes de la década de los ochenta aparecen referencias a la investigación-acción con otra denominación, será a partir del Seminario de Málaga (1984), «Métodos y técnicas de investigación-acción en la escuela», cuando en el Estado español surge el interés por este movimiento de investigación-acción. A juzgar por el número de seminarios (Granada, 1985; Murcia, 1986; Valencia, 1985), informes, cursos de doctorado (M. Bartolomé, 1985) y publicaciones que están apareciendo en el país, nos atrevemos a indicar que la investigación-acción ha adquirido un relieve peculiar y se está convirtiendo en el tema «estrella» de esta década en el área de la investigación educativa.

12.2 CONCEPTO Y CARACTERÍSTICAS

¿Qué se entiende por investigación-acción? ¿Cuál es su naturaleza y rasgos específicos? Contestar estas preguntas resulta tarea difícil para la mayoría de autores que han abordado el tema.

Escudero (1987, 21) advierte que la investigación-acción es una metodología de investigación educativa difícil de codificar en cánones precisos que permitan, con rigor lógico, acotar su conceptualización.

La investigación-acción se caracteriza por su naturaleza ambigua y heterogénea; admite variedad de usos e interpretaciones y carece de criterios claros y concretos para delimitar la gran variedad de orientaciones metodológicas que la reclaman para sí (Pini, 1981).

En los textos educativos aparecen diferentes expresiones, como *investigación en _el aula, el maestro investigador, investigación colaborativa, investigación participativa,* que designan modelos de investigación con cierta especificidad, pero que se consideran o como expresiones equivalentes a investigación-acción o como variantes de la misma.

La mayoría de los textos conceptualizan la especificidad de la investigación-acción como una metodología con algún rasgo particular relacionado con el rol del investigador, el fin de la investigación o el contexto social de la misma. Otros autores ven la especificidad de la investigación-acción como oposición al modelo tradicional de investigación educativa, y definida para ellos consiste en delimitar las fronteras con la investigación científica.

. El concepto de investigación acción ha ido cambiando con el tiempo y, especialmente su significado, aparece vinculado al contexto sociocultural que envuelve al investigador. Las definiciones de investigación-acción constituyen un amplio espectro

que va desde los que la consideran como un tipo de investigación aplicada (Bogdan y Billen, 1982) hasta los que quieren atribuirle el rango de paradigma singular y alternativo al modelo tradicional (Moser,1975; Escudero, 1987).

Con el tiempo el concepto de investigación-acción se flexibiliza para dar cabida a nuevas experiencias de investigación que afloran en el campo de la investigación educativa, llegando a considerarse en la actualidad como «palabra-paraguas» (Van Trier, 1980) o metaconcepto (Goyette y Lessard-Hébert, 1985), que engloba una amplia gama- de enfoques y estrategias de investigación cuya especificidad se caracteriza por ser una alternativa al modelo tradicional. Algunos autores ven en la investigación acción un modelo ideal al que aspiran algunas corrientes sociocríticas, y no tanto un modelo de investigación viable, cuestionando su implantación en li praxis educativa.

Veamos algunas definiciones. Corey (1949, 148) define la investigación-acción como «el tipo de investigación que se lleva a término en situaciones escolares y es diseñada para ayudar a la gente que allí trabaja a comprender si está actuando correcta o incorrectamente». Según Escudero (1987), Corey consideró la investigación-acción como un medio para desarrollar la capacidad de resolver problemas por parte del profesor, y como una metodología para elaborar el currículum y formar al profesorado.

Elliott (1981, 1) define la investigación-acción como «un estudio de una situación. social con el fin de mejorar la calidad de la acción dentro de la misma». La caracteriza como una reflexión sobre las acciones humanas y las situaciones sociales vividas por los profesores, que tiene como objetivo ampliar la comprensión (diagnóstico) del profesor de sus problemas prácticos. Las acciones van encaminadas a modificar la situación una vez que se logre una comprensión más profunda de los problemas.

Con Kemmis (1984, 1) la investigación-acción no sólo se constituye como ciencia práctica y moral, sino también como ciencia crítica. Para este autor la investigación-acción es «una forma de indagación autorreflexiva de los participantes (maestros, estudiantes o directores, por ejemplo) en situaciones sociales (incluyendo las educativas) para mejorar la racionalidad y justicia de: *a)* sus propias prácticas sociales. o educativas; *b)* la comprensión de tales prácticas, y e) las situaciones (e instituciones) en que estas prácticas se realizan (aulas o escuelas por ejemplo).

Finalmente, para Ebbutt (1983) la investigación-acción es un estudio sistemático orientado a mejorar la práctica educativa por grupos de sujetos implicados a través de sus propias acciones prácticas, y de reflexión sobre los efectos de tales acciones.

Kemmis y MacTaggart (1988, 30-34) señalan como puntos clave de la investigación-acción los siguientes:

1) Se propone mejorar la educación mediante su cambio, y aprender a partir de las consecuencias de los cambios.

2) Es participativa; las personas trabajan por la mejora de sus propias prácticas.

3) La investigación sigue una espiral introspectiva: una espiral de ciclos de planificación, acción, observación y reflexión.

4) Es colaborativa: se realiza en grupo por las personas implicadas.

5) Crea comunidades autocríticas de personas que participan y colaboran en todas las fases del proceso de investigación.

6) 6) Es un proceso sistemático de aprendizaje, orientado a la praxis (acción críticamente informada y comprometida).

7) Induce a teorizar sobre la práctica.
8)Exige que las prácticas, las ideas y las suposiciones sean sometidas a prueba.
9) Concibe de modo amplio y flexible aquello que puede constituir pruebas; implica registrar, recopilar, analizar nuestros propios juicios, reacciones e impresiones en tomo a lo que ocurre.
10) Exige llevar un diario personal en el que se registran nuestras reflexiones.
11) Es un proceso político porque implica cambios que afectan a las personas.
12) Implica la realización de análisis críticos de las situaciones.
13) Procede progresivamente a cambios más amplios.
14) Empieza con pequeños ciclos de planificación, acción, observación y reflexión, avanzando hacia problemas de más envergadura.
15) Empieza por pequeños grupos de colaboradores, expandiéndose gradualmente a un número mayor de personas.
16) Permite crear registros de nuestras mejoras (actividades, prácticas, lenguaje, discurso, relaciones, formas de organización).
17) Permite dar una justificación razonada de nuestra labor educativa mediante una argumentación desarrollada, comprobada y examinada críticamente a favor de lo que hacemos.

Para señalar los rasgos que configuran la investigación-acción podemos hacerlo a través de cómo considera algunos aspectos tales como:

a) Qué investigar. Existe alta coincidencia en asumir como foco de estudio la práctica educativa, tal como aparece contextualizada. Diagnostica los problemas en su contexto específico, intentando resolverlo. Predomina la práctica sobre la teoría. Suele centrarse en los problemas prácticos de los profesores.

b) Quiénes realizan la investigación. Asume la idea de los prácticos e investigadores trabajan en un contexto de colaboración, participación y democracia.

e) Cómo investigar. Suelen usar las estrategias de corte etnográfico/cualitativo: diarios, entrevistas, observación participante, notas de campo, procesos de triangulación y negociación.

d) Para qué investigar. Es uno de los rasgos más críticos y distintivos de la investigación-acción: contribuir a la resolución de problemas, cambiar y mejorar la práctica educativa

e) Naturaleza del objeto de investigación. La naturaleza social y humana de los fenómenos educativos; elude considerar la educación como «objeto natural». Es la l'!.áctica educativa la que se constituye en praxis informada.

/) Bases epistemológicas. Se adscribe a los postulados de la investigación interpretativa y crítica, se aparta de los principios del positivismo lógico.

g) Proceso. Consiste en una espiral de ciclos organizados en base a acciones planificadas y reflexiones críticas sobre las mismas.

Cohen y Manion (1985) agrupan los propósitos de la investigación-acción en cinco amplias categorías:

1) Es un medio de resolver problemas diagnosticados en situaciones específicas, o de mejorar una serie de circunstancias.

2) Es un medio de formación permanente.

3) Es un modo de insertar nuevos enfoques o innovaciones en la enseñanza y aprendizaje, en un sistema que de por sí inhibe la innovación y el cambio.

4)Es un medio de mejorar las comunicaciones entre los prácticos e investigadores.

5) Aunque sin el rigor de la investigación científica, aporta un enfoque alternativo preferible al sistema subjetivo e impresionista de resolver los problemas en el aula.

12.3 TIPOS DE INVESTIGACIÓN-ACCIÓN

La revisión de la literatura sobre este tema nos lleva a considerar dos líneas tipológicas: una basada en el modelo lewiniano (por ejemplo, Chein Y otros, 1948; Barbier, 1977; Werdelin, 1979) y otra en la escuela inglesa (P. Holy, 1984).

a) La tipología lewinianase basa en los objetivos de la investigación y señala cuatro modalidades:

- Investigación-acción diagnóstica: Los investigadores recogen datos, los interpretan, establecen un diagnóstico y recomiendan unas medidas de acción.
- Investigación-acción participativa. Implica a los miembros de la comunidad en el proyecto de investigación, considerándolos como agentes del proceso de investigación.
- Investigación-acción empírica. Estudia un problema social mediante una acción que supone un cambio, y valora los efectos producidos. Todo ello de la manera más sistemática posible.
- Investigación acción experimental. Difiere de la anterior en que la evaluación de los efectos del cambio se realizan a partir de un diseño experimental o cuasiexperimental.

b) La tipología de la escuela inglesa (Holly, 1984) establece tres amplias modalidades de investigación acción: técnica, práctica y crítica.

- Investigación-acción técnica. Su propósito es hacer más eficaz la práctica educativa y el perfeccionamiento del profesor, mediante la participación en programas de trabajo diseñados por un experto o equipo, en los que aparecen preestablecidos los propósitos el desarrollo metodológico que hay que seguir.
- Investigación-acción práctica. Confiere un protagonismo activo y autónomo a los profesores, siendo éstos quienes seleccionan los problemas de investigación y llevan el control del proyecto. Puede reclamarse la asistencia de un investigador externo, de otro colega o, en general, de un «amigo crítico».
- Investigación-acción crítica. Incorpora las ideas de la teoría crítica. Se centra en la praxis educativa, intentando profundizar en la emancipación de los profesores (sus propósitos, prácticas rutinarias, creencias), a la vez que trata de vincular su acción a las coordenadas sociocontextuales en las que se desenvuelve, así como la ampliación del cambio a ámbitos sociales.

12.4 PROCESO Y MODELOS DE INVESTIGACIÓN-ACCIÓN

12.5 El procesó de la investigación-acción se puede concebir como una espiral de ciclos constituidos por varios pasos o momentos. Su naturaleza flexible permite un permanente *feedback* entre cada una dé las fases o pasos del ciclo. El proceso se inicia con una «idea general» sobre las necesidades de mejorar o cambiar algún aspecto problemático de la práctica; a continuación se planifican los pasos y estrategias que hay que realizar; se lleva a cabo la acción, y termina el ciclo con la evaluación de los efectos de dicha acción, para volver a replantear el ciclo. Su carácter cíclico implica un «vaivén» (espiral dialéctica) entre la acción (praxis) y la reflexión (teoría), de manera que ambos momentos quedan integrados y se complementan.

Cohen y Manion (1985) señalan las ocho etapas y procedimientos que se pueden seguir en un diseño de investigación-acción. Ofrecen un marco básico, flexible, a modo de ilustración, que será necesario adaptar y contextualizar en cada caso particular:

1) Comprender la identificación, formulación y evaluación de un problema percibido como crítico en una situación de enseñanza. *Problema* tiene aquí un sentido amplio; puede referirse a la necesidad de introducir una innovación en el currículo.

2) Es la etapa de la discusión y negociación entre las partes implicadas (maestros, investigadores, consejeros, administradores, etc.) que culmina en una propuesta de proyecto de investigación.

3) Abarca la revisión de la literatura referente al tema de estudio: objetivos, procedimientos, resultados, etc.

4) Replanteamiento o redefinición del problema inicial a la luz de la revisión de la literatura.

5) Incluye la selección de los procedimientos de investigación: muestra, elección de materiales, asignación de recursos y tareas, organización del *staff*, etc.

6) Se eligen los procedimientos de evaluación que se usarán teniendo en cuenta que la evaluación en este contexto será continua,

7) En esta etapa se implementa el proyecto.

8) Se interpretan los datos; se infieren las consecuencias y se evalúa el proyecto a la luz de los criterios establecidos.

Para Escudero (1987, 2) el proceso de la investigación acción podría articularse en tomo a estas fases:

a) Identificación de un problema, dificultad o tema de investigación (análisis de la propia realidad para comprender mejor cómo y por qué ocurre).

b) Elaborar un plan de acción razonado (crear las condiciones para ponerlo en práctica).

c) Observar y controlar el curso, incidencias, consecuencias y resultados de su implementación, reflexionando críticamente sobre lo que sucedió, y elaborar una teoría situacional y personal de todo el proceso.

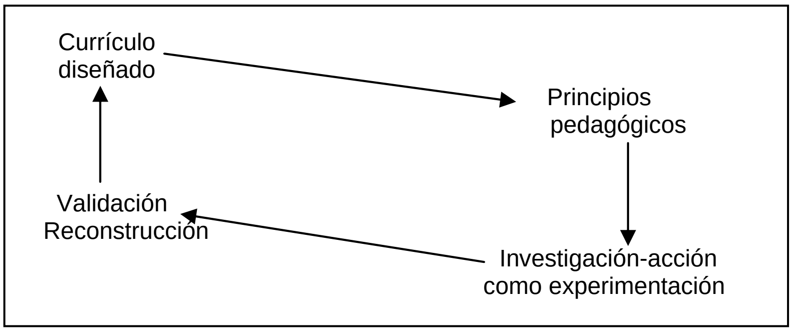


FIG. 12.1 Modelo investigación-acción (Escudero, 1987,24)

El currículo diseñado por los expertos tiene como soporte los principios que emanan de las disciplinas científicas y guían las prácticas educativas. Los prácticos experimentan si dichos principios y valores se cumplen en realidad, y a la luz de los resultados replantean el diseño curricular.

Veamos algunos modelos de investigación-acción. El formato y proceso son bastante similares con algunas variaciones que iremos comentando. Todos parten y se inspiran en el modelo lewiniano.

a) *El modelo lewiniano.* Lewin describió la investigación-acción como una espiral de ciclos. Cada ciclo se compone de una planificación, acción y evaluación del

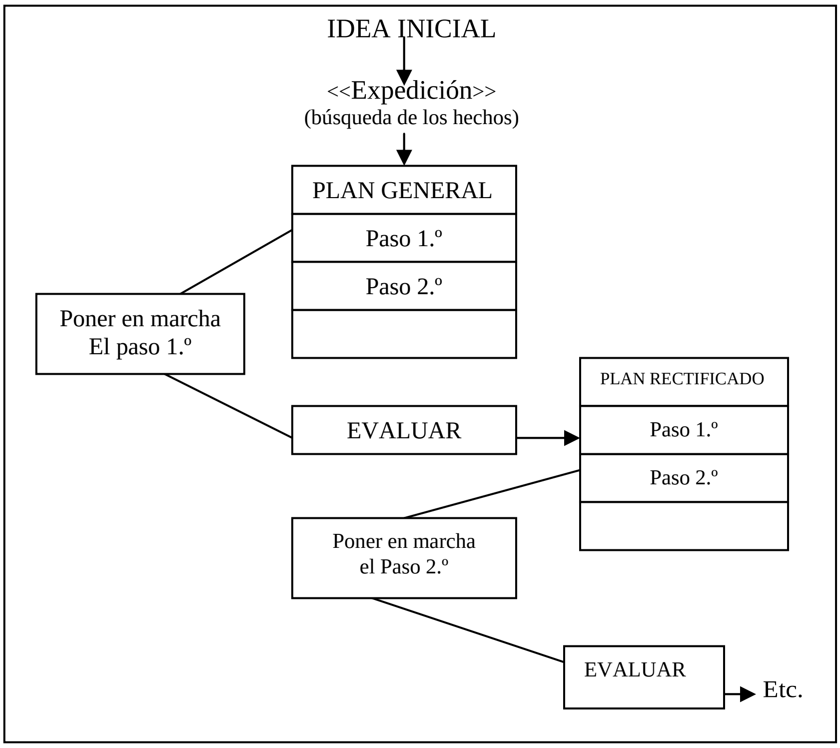


FIG. 12.2 El modelo de Lewin

resultado de la acción. Comienza con una «idea general» sobre un tema de interés sobre el que se elabora un plan de acción. Se hace un reconocimiento del plan, *sus* posibilidades y limitaciones, se lleva a cabo el primer paso de acción y se evalúa su resultado. El plan general es revisado a la luz de la información y se planifica el segundo paso de acción sobre la base del primero.

La figura 12.2 representa el proceso de investigación acción ideado por Lewin. b) *El modelo de Kemmis.* Este autor estructura el proceso de investigación-acción sobre dos ejes: uno estratégico, constituido por la acción y reflexión; y otro organizativo, constituido por la planificación y *la* observación (fig. 12.3). Ambas dimensiones están en continua interacción, de manera que se establece una dinámica que contribuye a salvar los obstáculos y a comprender los hechos que tienen lugar en la escuela.

	Reconstructiva	*Constructiva*
Discurso: entre participantes	*4. Reflexionar* Retrospectiva sobre la observación	*1. Planificar* Prospectiva para la acción
Práctica: en el contexto social	*3. Observar* Prospectiva para la reflexión	*2. Actuar* Retrospectiva guiada por la planificación

FIG. 12.3 Los «momentos» de la investigación-acción

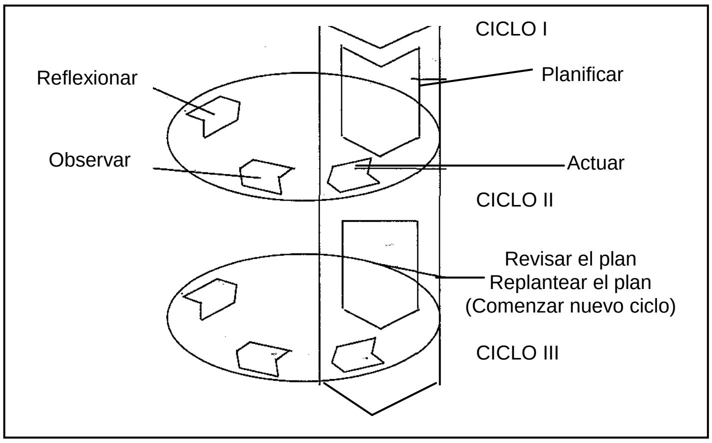


FIG. 12.4 Modelo de Kemmis, 1983

c) *El modelo de D. Ebbutt*: Defiende que la manera de entender el proceso de la investigación acción es concebido como una serie de ciclos sucesivos que proporcionen la posibilidad de retroalimentar la información entre y en cada uno de los ciclos del proceso. El proceso idealizado de investigación-acción educativa puede representarse mejor como muestra la figura 12.5.

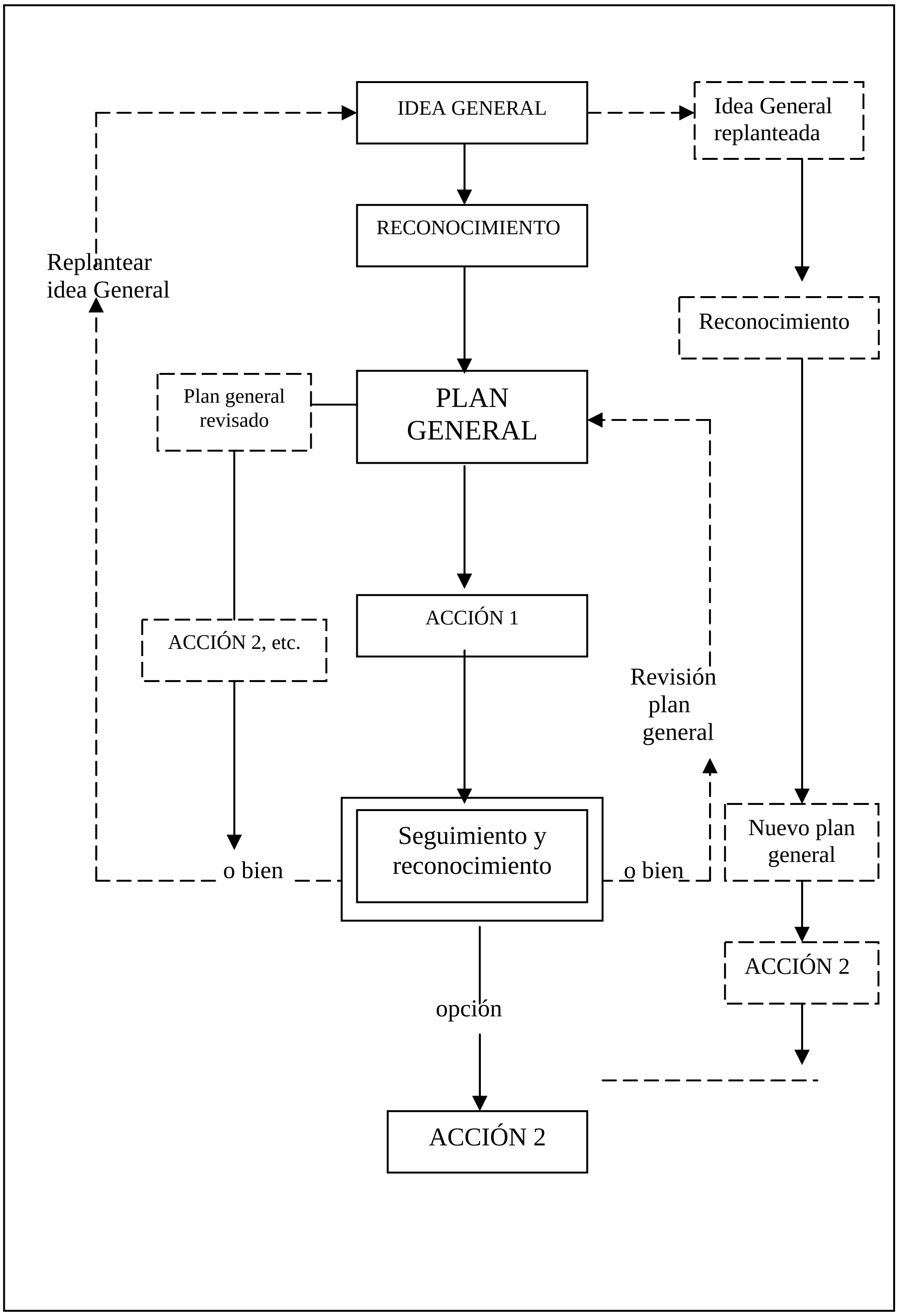


FIG. 12.5 Modelo de Ebbutt (1983)

d) *Modelo de J. Whitehead(1987).* Para que la investigación-acción tomara un carácter más real y personal para los prácticos, este autor reformula el ciclo de la investigación en frases que actúan de modelo general para tratar los problemas educativos prácticos de un modo sistemático. Éstas son:

1) Sentir o experimentar un problema. Siento un problema cuando algunos de mis valores educativos son negados en la práctica.

2) Imaginar la solución del problema. (Debería organizar las clases de forma que mis alumnos tengan que preguntar más.)

3) Implementar la solución imaginada. (Introduzco fichas de trabajo que lleve a mis alumnos a preguntar más.)

4) Evaluar la solución. (Mis alumnos participan más, pero se produce alboroto y desorden; están pendientes de mí cuando realizan las fichas.)

5) Modificar la práctica a la luz de los resultados. (Debo encontrar la solución para evitar el alboroto y conseguir que sean más independientes en sus tareas. [Mc Niff, 1988,38].)

12.5 INSTRUMENTOS Y TÉCNICAS

Los instrumentos y técnicas de la investigación-acción son los comunes de la investigación educativa, si bien recurre a aquellos que por su naturaleza son más propios del contexto de investigación que nos ocupa.

Este apartado recoge y enumera una serie de instrumentos genéricamente aplicables a la investigación-acción que se han utilizado en varios proyectos. Elliott (1986) propone las siguientes técnicas:

a) Diarios, que pueden ser tanto del investigador como de los alumnos o del profesor. Deberían contener observaciones, reflexiones, interpretaciones, hipótesis y explicaciones. Aportan información útil desde diferentes perspectivas.

b) Perfiles de secuenciación temporal. Proporcionan una visión de una situación o persona a través del tiempo. Debieran recoger una secuencia simultánea de tres componentes: actividad del profesor, actividad del alumno y materiales que ambas actividades precisan.

e) Análisis de documentos: programas, trabajos escolares, exámenes, actas de reuniones, circulares, periódicos escolares, etc., que permiten la comprensión con textual de un determinado ámbito educativo.

d) Fotografías y diapositivas (series de diapositivas) de momentos clave. Que sirven para ilustrar episodios concretos y de soporte visual de otras técnicas (notas de campo, entrevistas).

e) Grabaciones en audio y vídeo o transcripciones de las mismas. Permiten identificar áreas problemáticas y aportan evidencias sobre aspectos específicos de la enseñanza.

f) Participación de un observador externo. Puede realizar diferentes papeles, como facilitador de la investigación, observador participante, etc.

g) Entrevistas. Que pueden adoptar diferentes modalidades: entrevistas profesor/alumno, observador/alumno, alumno/alumno, cada una con diferente utilidad para la investigación.

h) Comentarios en vivo. Anotaciones de lo más relevante que acontece.

i) Estudio «en observación». Comentarios en vivo, continuos y sistemáticos, de uno de los participantes en la investigación.

j) Listas, cuestionarios e inventarios.

k) Triangulación. Técnica tanto de recogida de datos como de validación de información o evidencias. En el primer caso permite triangular fuentes de datos.

l) Informes analíticos. Especie de memorando que recoge las evidencias logradas en el período de investigación.

12.6 r.,A INVESTIGACIÓN P ARTICIP A TIV A

Investigación participativa *(participatory research)* es la expresión que adopta y promueve el movimiento de educación de adultos, ICAE (International Council for Adult Education, radicado en Taranta, Canadá, 1973), para configurar el enfoque participativo, activo, democrático ... de educación de adultos. Asimismo, bajo este epígrafe se engloba una modalidad de investigación que tiene lugar en los países del tercer mundo, con el objetivo de emancipación y movilización individuales mediante diálogo y participación en análisis críticos y acciones organizadas para mejorar su .. situación (Rosenwald, 1988, 20).

Autores como Werdelin (1979) y Hall (1981) ven en la investigación participativa una forma de investigación-acción social que surge como alternativa metodológica para canalizar las investigaciones donde las dimensiones participativa, democrática, liberadora, crítica y comprometida son nucleares, teniendo como propósito promover el «desarrollo comunitario». Cuando la investigación-acción se aplica a colectivos humanos en forma de grupos sociales, sobre todo grupos marginados, surge la investigación participativa.

CONCEPTO y CARACTF..RÍSTICA5....

Hall (1981, 7)define la investigación participativa como «actividad integrada que combina "la investigación social, el trabajo educativo y la acción». En esta definición se intregan las tres características que configuran la investigación participativa: como método de investigación social, como proceso educativo y Calla, medio para adoptar decisiones para el desarrollo (Werdelin, 1979, 10).

Para Le Boterf (1985) la investigación participativa es un proceso en el cual !QS, actores sociales no son considerados como simples objetos pasivos de investigación, que se transforman cada vez más en sujetos que conducen una investigación en colaboración con investigadores profesionales.

La investigación participativa reúne a investigadores y participantes en un diálogo que expande el conocimiento y la conciencia de ambos, al poder unir la acción para resolver problemas compartidos. El conocimiento puede expandirse con nueva información; la conciencia crítica de uno se puede mejorar con el diálogo; nuevas iniciativas y acciones se pueden inspirar en nuevos *insights* (Brown, cito por Rosenwald, 1988). Investigación y educación se consideran como momentos de un mismo proceso, es decir, la investigación participativa se transforma en quehacer de aprendizaje colectivo (Boris Yopo, 1981).

Las características que enumeramos más abajo son un sumario de varios autores que han abordado este tema (como Tandon, 1981; Le Boterf, 1986; Werdelin, 1978; ICAE, 1978). Las notas más relevantes de la investigación participativa serían estas:

a) Su carácter democrático. Una idea nuclear es la relativa a la toma de decisiones conjuntas para definir los problemas, diseñar la investigación, recoger datos, interpretar y valorar los resultados, etc.

b) La participación e implicación de la comunidad (grupo social, comunidad, escuela) es clave, variando en su naturaleza y extensión.

c) Los participantes asumen el control del proceso y de la toma de decisiones.

d) Es un proceso colectivo que conlleva una experiencia educativa. Participar implica adquirir conocimiento más objetivo de la situación, analizar con más precisión sus problemas, descubrir los recursos de que disponen, formular las acciones pertinentes.

e) El investigador es un participante comprometido que a la vez que conduce al grupo con imparcialidad aprende del proceso de investigación. Asume varias responsabilidades: puede ser el iniciador de la investigación, la persona recurso a quien los participantes pueden dirigirse para aconsejarse e informarse, o el profesor que entrena a los participantes a jugar sus papeles.

/) En la recogida de datos se puede utilizar la investigación tradicional, pero dando mucho peso a las posturas cualitativas e interpretativas y a la comunicación interpersonal.

g) Entre la investigación y la acción existe una interacción permanente. La acción es fuente de conocimiento y la investigación constituye en sí una acción transformadora.

h) La investigación participativa se realiza en situaciones naturales de los sujetos; evitando las artificiales o de laboratorio.

i) Entre los participantes en la investigación se crea una situación de interacción activa, de diálogo y negociación.

j) Se pone al servicio de grupos o categorías sociales más desfavorecidos.

k) La evaluación se realiza cooperativamente con los participantes en la comunidad o grupo estudiado.

FINES y OBJETIVOS

La investigación participativa sirve a varios fines u objetivos, de los que queremos señalar los siguientes:

a) El fin último de la investigación participativa es la trasformación de la realidad s9siél1y promoción del desarrollo comunitario.

b) Crear en los participantes autoconciencia de su realidad social y capacidad para tomar decisiones para mejorar1a.

e) Activar a los participantes, Capacitar a la gente para movilizar sus recursos humanos para la solución de los problemas sociales.

á) Concienciar, actuar, liberar son ideas que presiden la actuación de la investigación participativa.

e) El desarrollo de la sociedad. Puede contribuir a su desarrollo en más de un sentido: 1) desarrollando las capacidades de sus miembros; 2) analizando necesidades,

fines, demandas, problemas y oportunidades; 3) encontrando soluciones a algunos problemas; y 4) iniciando otras actividades sociales.

PROCESO DE LA INVESTIGACIÓN PARTICIPATIVA

La investigación participativa se adscribe a las corrientes metodológicas interpretativas y crítica, con las que comparte los fundamentos epistemológicos, metodológicos y estratégicos. Lo que no obsta para que su peculiar metodología permita que las decisiones se tomen democráticamente; que el problema se defina en función de una realidad concreta y compartida; que los grupos sociales elaboren la programación del estudio y las formas de encararlo, con sus modalidades de gestión y evaluación de un proyecto concreto de acción para el desarrollo tecnosocial de la propia comunidad (Boris Yopo, 1981).

La naturaleza de la investigación participativa precisa quela metodología se adapte a cada proyecto de investigación. Se trata de adaptar el proceso de investigación a las condiciones particulares (recursos, contexto sociopolítico, objetivos perseguidos, etc.) que la situación concreta requiere.

El principio fundamental y su punto de partida más radical es que el proceso de investigación se basa en un sistema de discusión, investigación y análisis, en el que los investigados son tanto más una parte del proceso como el investigador (Hall, 1977). Ilustramos la metodología de investigación participativa con el modelo propuesto por Le Boterf (1980).

1.ª fase. Montaje institucional y metodológico de la investigación participativa. Conlleva las tareas:

a) Discusión del proyecto con la población y sus representantes.
b) Definir el cuadro teórico: objetivos, conceptos, hipótesis, métodos. e) Delimitar la zona, población y problema que hay que estudiar.
d) Organización del proceso de investigación.
e) Seleccionar y formar investigadores o grupos de investigación.
f) Elaborar y aprobar el presupuesto.
g) Elaborar un calendario de acciones que hay que realizar.

2.ª fase. Estudio o diagnóstico previo de la zona y de la población. Incluye tres componentes:

a) Identificación de la estructura social de la población.
b) Descubrimiento de los problemas vividos por la población y su naturaleza. e) Censo de los datos socioeconómicos y tecnológicos.

3.ª fase. Análisis crítico de los problemas. Tres momentos:

a) Expresión del problema.
b) Cuestionamiento del mismo.
e) Reformulación del problema: descripción, explicación, estrategias de acciones posibles.

4.ª fase. Elaboración y aplicación de un plan de acción (incluyendo las acciones educativas) que contribuyan a la solución de los problemas. Comporta:

a) Las actividades educativas que permitan analizar los problemas y las situaciones vividas.
b) Las medidas que puedan mejorar la situación a nivel local.
e) Las acciones educativas que posibiliten aplicar dichas medidas.
d) Las acciones para poner en marcha las soluciones previstas a corto, medio y largo plazo, a nivel local o a escala más amplia.

El proceso de la investigación participativa no acaba con las cuatro fases descritas. El análisis crítico de la realidad, la realización de las acciones programadas, conducen al descubrimiento de los problemas, otras necesidades, otras dimensiones de la realidad. La acción es una fuente de conocimiento y nuevas hipótesis. El diagnóstico, el análisis crítico y la acción constituyen así tres momentos de un proceso permanente de estudio, reflexión y transformación de la realidad que se nutren mutuamente (Le Boterf, 1980).

CAMPOS DE APLICACIÓN

La investigación participativa se ha utilizado mayormente en el ámbito de la educación de adultos, lo que no quiere decir que no se haya utilizado o se pueda utilizar en otros ámbitos. Werdelin (1978) señala los siguientes:

1) Educación de adultos.
2) Educación no formal de adultos.
3) Programas de entrenamiento para pequeñas y medianas empresas.
4) Desarrollo organizacional
5) Desarrollo de programas escolares.
6) Proyectos interdisciplinares y participación de estudiantes.
7) Programas de desarrollo para la función gestora.
8) Estudios de proyectos.

12.7 LA INVESTIGACIÓN COLABORATIV A

La investigación colaborativa *(collaborative research),* expresión que se utiliza como sinónimo de *investigación-acción colaborativa* y de *investigación interactiva* es una modalidad de investigación-acción que surge en los Estados Unidos para designar un tipo de investigación cuyo rasgo básico sería «investigación en colaboración>>, «coinvestigación>>. Un equipo de personas (profesores, investigadores, alumnos, padres, etc.) investigan conjuntamente situaciones o problemas compartidos.

La colaboración, como concepto de aprendizaje e investigación, describe una forma peculiar de implicación participativa en los proyectos de investigación y otros programas de aprendizaje. Adopta el principio de trabajar con, no trabajar «sobre» los profesores y escuelas (Lieberman, 1986).

Investigación colaborativa significa implicación de investigadores, profesores, padres, administradores y estudiantes en proyectos comunes; lo que no significa que cada uno tenga el mismo rol en la toma de decisiones o *inputs* durante las etapas o fases del estudio. Los cambios de rol ocurren dependiendo de las" necesidades de la situación. La continuidad es prevista por los investigadores mediante la comunicación y una red de colaboración que establecen con los implicados en el estudio (Wallat, 1981).

Si bien la expresión *investigación colaborativa* es reciente, las ideas que la configuran hay que situadas en el trabajo de Corey, quien enfatiza el estudio cooperativo de los problemas por parte de los prácticos e investigadores, y en la obra de Lewin, quien había creado una serie de conceptos y definiciones relativos al estudio cooperativo, que Corey usó en el desarrollo de la investigación-acción (Jacullo-Noto, 1984). En este sentido se puede decir que la investigación colaborativa surge como desarrollo de la investigación-acción en el campo colaborativo. Los trabajos de Tikunoff y Ward (1982), Smulyan (1984), Lieberman (1986), Jacullo-Noto (1984), Oja y Pine (1983), etc., son algunos de los más representativos.

Smulyan (1984) señala tres etapas fundamentales en el origen y evolución de la investigación colaborativa: la primera coincide con la investigación-acción en sus orígenes y expansión; la segunda señala el declive de ésta y la aparición de otros enfoques de investigación, y la tercera recoge el nacimiento de la investigación colaborativa.

CONCEPTOS Y CARACTERÍSTICAS

La investigación colaborativa se presenta como un modelo alternativo de indagar la realidad educativa. Su definición pone el énfasis en el hecho de que investigadores y educadores coinvestigan. Trabajan conjuntamente en la planificación, implementación y análisis del proceso la investigación que se lleva a cabo para resolver los problemas inmediatos y prácticos de los maestros, compartiendo la responsabilidad en la toma de decisiones y en la realización de las tareas de investigación (Bartolomé, 1986,54).

Es conceptualizada como un proceso de indagación y teorización sobre las prácticas profesionales de los prácticos y las teorías que guían estas prácticas. Es una forma de investigación que capacita a la comunidad de educadores para mejorar lo que hacen y comprender mejor su práctica. Es una forma de investigación que genera información y teorías no como productos para ser sintetizados y expuestos, aislados de la fuente práctica, sino como *insights* contextualizados y personalizados en el propio desarrollo personal, como un educador, un estudiante, un trabajador social, un político ... (Henry, 1986,86).

Para JacuIlo-Noto (1984) consiste en una estrategia que implica a investigadores y prácticos en un proceso de investigación y desarrollo interactivo. El trabajo en colaboración tiene un Claro objetivo: trabajar juntos sobre aspectos profesionales compartidos; es una actividad de investigación y aprendizaje para los profesores interesados en indagar conjuntamente para resolver problemas de enseñanza/aprendizaje confrontándolos con su práctica educativa.

9j[a] y Pine (1981, 9-10) señalan que la investigación-acción colaborativa se caracteriza por varios elementos:

1) Los problemas de investigación son mutuamente definidos por prácticos e investigadores.

2) Los investigadores y los profesores colaboran en la búsqueda de soluciones para los problemas del práctico.

3) Los resultados de la investigación son utilizados y modificados en la solución de los problemas.

4) Los prácticos desarrollan competencias, habilidades y conocimiento de investigación, en tanto que los investigadores se reeducan utilizando metodologías naturalistas y etnográficas de investigación.

5) Los prácticos, como resultado de su participación en el proceso de adaptación, son más capaces de resolver sus propios problemas y renovarse profesionalmente. 6) Prácticos e investigadores son coautores de los informes de investigación.

Jacullo- Noto (1984) añade además: la interacción y concurrencia de investigación y desarrollo ha de tener lugar a lo largo de todo el proceso y se ha de mantener la integridad natural del contexto: la clase no ha de verse alterada a lo largo del desarrollo del proyecto.

Existen tres supuestos que subyacen a este enfoque:

1) Paridad en la toma de decisiones entre investigadores, formadores y prácticos.
2) Respeto a la propia perspectiva de cada componente del equipo.
3) Responsabilidad compartida entre cada participante en la investigación colaborativa y proceso de desarrollo teórico (Mergendoller, 1979, cit. por Smulyan, 1984).

PROCESO DE LA INVESTIGACIÓN COLABORATIVA

Uno de los requisitos previos al proceso de investigación colaborativa es que los miembros del equipo aprendan a colaborar. ¿Qué es colaborar?, ¿qué implica?, ¿cuál es su coste?, ¿cuáles son sus riesgos?, ¿cuáles sus beneficios?, son preguntas a las que los participantes tienen que dar respuesta. La colaboración no es algo connatural; es un proceso sofisticado que debe ser enseñado y aprendido deliberadamente.

Para abordar el tema del proceso de investigación seguimos el modelo citado por M. Bartolomé (1986, 64), que contiene los pasos siguientes:

a) A partir de un problema concreto, vivido por los profesores, se intenta sistematizar de alguna forma la dificultad surgida a partir de recogidas de datos, dentro de un contexto determinado.

b) El equipo discute y elabora categorías básicas, que permitan sintetizar y comparar los datos que se van obteniendo a lo largo del tiempo.

e) Se acumula evidencia empírica sobre la cuestión estudiada por la utilización de diversos procedimientos de recogida de datos.

el) La interpretación de los resultados en el seno del equipo permite enriquecer la visión del problema al tiempo que puede significar procesos de cambio en la mentalidad y forma de actuar en los participantes ..

e) Se establecen regularidades y relaciones entre los datos observados.

f) Sucesivos exámenes permiten afianzar las relaciones descubiertas (que pueden generar y expresarse en forma de hipótesis).

g) Finalmente, puede obtenerse cierta estructura de generalización, dentro de ese contexto, que posibilite la elaboración de teorías al tiempo que facilite propuestas de solución y cambio en el ámbito educativo.

En la tabla 12.1 se recogen las distintas etapas de dos procesos de investigación colaborativa propuestos por Tikunoff y otros (1979) y por Smulyan (1983).

TABLA 12.1 Procesos de investigación co1aborativa (cit. Bartolomé, 1986,65)

Etapas del proceso de investigación colaborativa	
1. Identificación de un aspecto que ha de ser estudiado. 2. Selección de estrategias de investigación. 3. Naturaleza de los datos que han de recogerse. 4. Procedimientos de recogida de datos. 5. Preparación del diseño de investigación. 6. Recogida de datos. 7. Análisis de datos. 8. Desarrollo de conclusiones para realizar el informe.	1. Identificación de un problema. 2. Identificación de una cuestión que hay que investigar dentro de ese problema. 3. Discusión de la metodología: a) Procedimientos de recogida de datos. b) Naturaleza de los datos que han de ser recogidos. 4. Preparación del diseño de *investigación.* 5. Recogida de datos. 6. Análisis de datos. 7. Presentación de conclusiones.
(Tikunoff y otros, 1979, 148)	(Smulyan, en Oja y Pine, 1983, 425)

CONDICIONES PARA REALIZAR INVESTIGACIÓN COLABORATIVA

El éxito de la investigación colaborativa depende de que se den una serie de condiciones. A modo de resumen indicamos las siguientes (M. Bartolomé, 1986,67):

a) Clima escolar. Se precisa un clima de respeto y libertad hacia la actividad investigadora de los profesores. Apoyo institucional y reconocimiento y legitimación por parte de la dirección del centro.

b) Disponer de recursos. Apoyo por parte de la Administración. Asistencia técnica y recursos de investigación: medios audiovisuales, subvenciones, servicio de consultas, libros, etc.

e) Formación. Una de las ideas reiterativas de la investigación colaborativa es que el éxito depende de las características y habilidades del equipo investigador. La formación debería tener una doble orientación: hacia la adquisición de competencias y hacia el desarrollo de la maduración personal y grupal.

d) La relación con el proceso. Las condiciones más citadas son: claridad en las metas, tiempo necesario y poder establecer una dinámica grupal que favorezca el desarrollo paulatino del *staff,* entendido éste como el «intento sistemático de alterar las prácticas profesionales, creencias y conceptos del personal de la escuela ante un fIn articulado» (Griffin, 1983).

Smulyan (1984, 16) señala cinco aspectos como condiciones de éxito de un proyecto de investigación c01aborativa: las características del profesor, la organización de la escuela, el clima escolar, los recursos disponibles y la estructura del proyecto de investigación.

Muchos de los problemas implicados en la realización de la investigación-acción colaborativa proceden del mismo elemento que contribuye a su valor: su naturaleza colaborativa. Smulyan (1984) señala los siguientes:

a) El inicio del proyecto de investigación; los prácticos no suelen estar tan dispuestos como los investigadores, desconfían de la habilidad de los investigadores para resolver sus problemas.

b) La diversidad de intereses de los participantes; los prácticos buscan encontrar vías de mejora de su enseñanza o escuela, mientras que los investigadores buscan generalizaciones que puedan compartir con la comunidad educativa.

e) Los problemas que surgen en los procesos de colaboración, que se dan entre el inicio del proyecto y la producción de resultados: la interpretación de cada uno del significado del proceso, la diversidad de lenguajes, el rol que cada uno asume, etc. (Smulyan, 1984, 19-22).

En las últimas décadas han aparecido nuevas concepciones y metodologías que han tenido una profunda repercusión en la transformación de los planteamientos de la investigación educativa. Esta obra ofrece una visión general actualizada de dichas innovaciones. Recoge las nuevas tendencias en investigación etnográfica, evaluativa e investigación acción que permiten acercarse, penetrar, conocer e interpretar la realidad educativa desde diferentes perspectivas.

Investigación educativa es un manual básico de referencia que compagina el rigor y sistematización de los contenidos con un lenguaje comprensible.

Es una obra de gran utilidad para estudiantes universitarios y también para los profesionales que deseen profundizar en el conocimiento de los procesos y metodologías de la investigación educativa.

Los autores son profesores titulares de reconocido prestigio. Justo Arnal ejerce su labor docente en el Departamento de Pedagogía y Didáctica (Universidad Autónoma de Barcelona), y Antonio Latorre y Delio del Rincón, en el Departamento de Métodos de Investigación y Diagnóstico en Educación (Universidad de Barcelona).

9 788433 537256